DOCUMENT RESUME

ED 219 444 TM 820 468

AUTHOR	Alkin, Marvin C.; And Others
TITLE	Title I Evaluation: Utility and Factors Influencing Use.
INSTITUTION	Educational Evaluation Associates, Los Angeles, Calif.
SPONS AGENCY	Office of Program Evaluation (ED), Washington, DC.
PUB DATE	[82]
NOTE	193p.
EDRS PRICE	MF01/PC08 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; Federal Programs; Information Utilization; *Program Evaluation; *School Districts; *State Departments of Education
IDENTIFIERS	Elementary Secondary Education Act Title I; *Evaluation Utilization; *Title I Evaluation and Reporting System

ABSTRACT

This study investigated the utility of Title I evaluation, and in particular, the Title I Evaluation and Reporting System (TIERS). To examine utility, instances of evaluation use in state (SEA) and local education agencies (LEA) were identified, and the factors associated with that use were studied. Title I evaluation was used by SEAs to: insure local compliance with state and federal requirements; assist in recognition of problems that required programmatic intervention at the local level; provide a basis for recommending solutions to those problems; and act as input to educational decisions under consideration at the state level. At the LEA level, Title I evaluation data were used by a variety of personnel--including school boards and other external agencies, district administrative staffs, school principals, and individual school personnel to: modify attitudes; recognize or identify problems; prompt administrative or organizational action; and induce curricular or instructional change. The factors most responsible for the use of evaluation information were: the reputation of the evaluator; the evaluator's commitment to evaluation use; the interest of decisionmakers and the community; the extent to which the evaluation focuses on local needs; the degree to which evaluation is presented in graphic, non-technical form; and the development of procedures that assist decisionmakers to use evaluation information. (Author/BW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED219444

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. C. Alkin

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TITLE I EVALUATION:
UTILITY AND FACTORS INFLUENCING USE

Presented to the
U.S. Department of Education,
Office of Program Evaluation

TM 820468

EEA EDUCATIONAL EVALUATION ASSOCIATES

TITLE I EVALUATION: UTILITY AND FACTORS INFLUENCING USE

(Presented to the U.S. Department of Education, Office of Program Evaluation)

by

Marvin C. Alkin

Brian M. Stecher

Frederica L. Geiger

EEA

EDUCATIONAL EVALUATION ASSOCIATES
9230 Jellico, Northridge, California 91324
(213) 993-8070

ACKNOWLEDGEMENTS

We would like to acknowledge the contributions of Larry Braskamp, Richard Daillak, Jean King, and Richard Williams, who conducted extensive interviews with LEA personnel and prepared the case studies which make up Appendices B-E. We would also like to thank the TAC directors for providing insight into the Title I activities in their regions and suggesting exemplary SEAs and LEAs. The cooperation of the SEA and LEA personnel we contacted was outstanding. To maintain anonymity, we cannot mention names, but we want to acknowledge their assistance in making this study possible. In particular, we would like to thank "Jack Brown," "Harold Green," "Sally Lopenia," "Bess Foley," and "Bill Leonard," the pseudonymous evaluators in the case studies. Finally, we want to recognize the editorial assistance provided by Gretchen Flesher and Jack McCurdy.

EXECUTIVE SUMMARY

This study investigated the utility of Title I evaluation, and in particular, the Title I Evaluation and Reporting System (TIERS). Since utility means the potential for utilization, to examine utility we sought instances of evaluation use and examined the factors associated with that use. Consequently, the study focused on state and local educational agencies that were believed to have utilized Title I evaluation information.

This study found evidence of utilization at both SEA and LEA levels. Title I evaluation was used by SEAs to:

- insure local compliance with state and federal requirements
- assist in recognition of problems that required programmatic intervention at the local level
- provide a basis for recommending solutions to those problems
- act as input to educational decisions under consideration at the state level.

At the LEA level, this study found substantial Title I evaluation use. Information from Title I evaluation data was used by a variety of personnel--including school boards and other external agencies, district administrative staffs, school principals, and individual school personnel to:

- modify attitudes
- recognize or identify problems
- prompt administrative or organizational action
- induce curricular or instructional change.

From the in-depth case studies we were able to identify the factors most responsible for the use of evaluation information. The most important factors appeared to be:

- the reputation of the evaluator
- the evaluator's commitment to evaluation use
- the interest of decisionmakers and the community
- the extent to which the evaluation focuses on local needs
- the degree to which evaluation is presented in graphic, non-technical form
- the development of procedures that assist decisionmakers to use evaluation information.

Results of the study led us to suggest several recommendations for improving evaluation utility in the future:

- provide technical assistance in reformatting TIERS-like results to meet LEA needs
- encourage SEA and LEA evaluation units to design evaluation strategies focused on local decision needs
- develop procedures which show decisionmakers how to use evaluation data
- train evaluators to use interpersonal skills to increase the effectiveness of their roles.

TABLE OF CONTENTS

	Page
Executive Summary	ES-1
Chapter 1 - The Utility of Title I Evaluation	1
Chapter 2 - Review of Research Literature and Technical Documentation	9
Chapter 3 - Title I Evaluation Activities and Use - SEA Level	39
Chapter 4 - Title I Evaluation Activities and Use - LEA Level	57
Chapter 5 - Factors Affecting Utilization in the Case Studies	81
Chapter 6 - Conclusions and Suggestions for Future Action	95
Bibliography	99
Appendix A - Data Collection Instruments	101
Appendix B - Middle City, Midwest: A Study of Title I Evaluation Use by Larry Braskamp	109
Appendix C - Big City, East: A Study of Title I Evaluation Use by Richard Daillak	129
Appendix D - Small Town, South: A Study of Title I Evaluation Use by Jean King	151
Appendix E - Suburb, West: A Study of Title I Evaluation Use by Richard Williams	175

LIST OF TABLES

Table	Page
1. Instances of Evaluation Use in Exemplary Districts Number of Incidents (Unduplicated)	60

CHAPTER 1

THE UTILITY OF TITLE I EVALUATION

There is increasing concern over the degree to which program evaluations have practical utility for educational decisionmakers. Utility in the broad sense refers not only to actual utilization but also to the potential for utilization. In this study, we examine instances in which utilization has occurred at the state educational agency (SEA) and local educational agency (LEA) levels, and we examine the contextual factors that have contributed to this utilization. Our intent is not to measure systematically and exhaustively the level of utilization. Rather, by identifying certain instances of utilization, we attempted to determine whether the Title I evaluation system has utility. Secondarily, we hoped to uncover some of the contextual variables that affect utility.

We used a broad definition of utilization in this study. The review of research contained in Chapter 2 suggests that in studies of Title I evaluation use, as in other utilization studies, restricting the definition of utilization to identifiable impact on specific actions ignores a great many important uses of evaluations. The perspective on evaluation use this research team subscribes to was described by the project director in an earlier work (Alkin, Daillak and White, 1979). Utilization may include gradual incremental influences on attitudes and judgments throughout the entire course of decision formulation. This conception may be contrasted with the narrower view which considers utilization in clear-cut terms at the actual point of decision.

Methodology

Obtaining a Sample

Our strategy was to study SEAs and LEAs with exemplary Title I evaluation utilization. Therefore, we began our search for exemplary SEAs and LEAs by contacting the Title I Technical Assistance Centers (TAC). This network of 10 Title I funded agencies throughout the country was designed to help SEAs and LEAs conduct Title I evaluations. The TACs have worked in all 50 states as well as several territories. They have consulted with Title I evaluation units in all State Departments of Education, and they have developed materials, presented workshops and seminars, and given technical assistance in a vast number of LEAs throughout the country. As a result, the TACs seemed to be a valuable resource to assist in the selection of SEAs and LEAs for further investigation.

The principal investigator telephoned the director of each of the 10 TACs to discuss the nature of the Title I evaluation activities taking place within that region.[1] We described our definition of evaluation utilization to each TAC director, and they discussed the nature of Title I evaluation activities conducted in each state in the region as well as the utility of the evaluation information at the state level.

Often the principal investigator had several conversations with a TAC director extending over a three or four day period. This allowed the TAC director to consult with staff to obtain specific information

---

[1] In one instance, the TAC director was on extended leave and we spoke with the assistant director.

about particular instances or types of utilization that might occur within an LEA or an SEA.

TAC directors were asked to recommend SEAs which they believed demonstrated a high level of Title I evaluation use. Based on these nominations, 12 state educational agencies were selected for interviewing. We divided the country into four geographic regions (East, West, Midwest and South) and interviewed three state educational agencies within each region.

SEA Interview Procedure

The data collection could, due to time constraints, last only five weeks. The compressed data collection timetable did not allow us to develop formal instruments and obtain clearances for their use. Thus, out of necessity, we adopted open-ended interviewing procedures.[2] The discussions with SEA personnel all followed a similar agenda, but the specific questions that were asked, the clarifications that were requested, etc. varied from conversation to conversation in response to the comments of the SEA personnel.

---

[2] The choice of research strategy was not, however, dictated solely by necessity. We have found in our previous research that most subjects have a very limited view of the meaning of utilization. The common tendency to try to recall a specific instance in which evaluation information was presented and a particular decision was made. This type of utilization rarely occurs. Indeed, the impact of evaluation is incremental and occurs over a considerable period of time. (See Alkin et al., 1979, and Patton, 1978.) Thus, to really assess the impact of evaluation at the LEA level one must adopt a research strategy that is responsive to subtle evaluation influences. Naturalistic inquiry offers such an alternative, and a key aspect of many naturalistic approaches is the open-ended character of the interchange. Thus, we would probably have adopted open-ended interviews for this study under any circumstances.

The following topics represent the kinds of concerns addressed in each interview, though they were discussed differently in almost every instance. The general agenda was:

1. Introduction -- indicating referral from the TAC, and the sponsorship of the U.S. Department of Education for this study.

2. Comments on utilization -- indicating the focus of the study and its orientation to evaluation utilization--viewing evaluation as one of a number of potential inputs into the decision process, contributing to decisions as well as changes in attitudes, procedures, etc.

3. General overview of SEA evaluation and utilization -- attempting to get respondent to talk about the nature of Title I evaluation in that particular SEA, including some thoughts on how it is utilized.

4. Problem recognition -- requesting that the respondent consider and discuss situations in which evaluation information contributed to the recognition of a problem (but not necessarily to the solution). In this category we wanted to identify the nature of the problems that were recognized, and the kind of Title I evaluation data that contributed to recognition.

5. Problem solution: administration -- expressing an interest in the extent to which statewide Title I evaluation data were a source of information related to modifications in administrative procedures at the state level.

6. Problem solution: instructional improvement -- expressing an interest in determining whether there were instructional changes or other substantive changes in the Title I program that had been impacted by statewide Title I evaluation data.

7. Factors affecting utilization -- seeking to determine contextual, political, or personal factors that seemed to affect statewide evaluation utilization. (There is, of course, a body of literature related to factors affecting evaluation utilization but only one national study related to Title I--and that was conducted prior to the introduction of the Title I Evaluation Reporting System (TIERS). Moreover, most of the rest of the widely-known literature on educational evaluation utilization was not focused at the SEA level.)

Typically, we interviewed one or two people within each SEA. We began with the official who was designated Title I contact person, and

in a number of cases, we were referred to another individual for additional information. In many instances the respondent chose to gather materials, check files, discuss issues with colleagues, etcetera in order to clarify certain points that had been raised.

LEA Selection

Both TACs and SEAs were asked to recommend LEAs that had high levels of Title I evaluation use. TAC directors suggested LEAs that might possibly be included in the study, or identified particular states they believed might have exemplary LEAs to include. LEA recommendations were also solicited from SEA personnel during the state interviews. In most cases the actual selection of LEAs was not made until discussions had been held with the Title I evaluation contact person in the appropriate SEAs. Thus, by necessity, the final selection of LEAs occurred after the SEA interviews were concluded.

We selected a group of 20 LEAs distributed equally among the four geographic regions for initial telephone interviews. The 20 selections were based on a number of factors including the likelihood of finding relevant information on evaluation utilization, the ease of further data collection, and district cooperation. In one or two instances, the recommendation of the TACs, as well as our prior knowledge of the LEAs, were so unequivocable that we could select a local district without further inquiry. However, in most cases our selections were based on subsequent consultation with representatives of the state agencies.

The 20 LEAs we selected were not all located within the 12 states we interviewed at the state level. Our own knowledge of some situations (as well as the comments of TAC directors) led us to believe that it was productive to interview a state at the SEA level, but it would be less relevant to select an LEA within that state. We also suspected that there were states in which there was far more utilization activity at the LEA level than at the SEA level.

LEA Interview Procedure

In order to complete the data collection as efficiently as possible, we obtained the assistance of a trusted colleague in each geographic region. There were two major criteria for the selection of these researchers: knowledge of evaluation utilization and previous publication in the area of evaluation utilization. In addition, each of the regional interviewers we selected subscribed to the "alternative" (or broader) perspective of utilization prescribed by the project director in an earlier work (Alkin, Daillak and White, 1979). The regional interviewers were: East, Richard Daillak (St. John's University); Midwest, Larry Braskamp (University of Illinois); South, Jean King (Tulane University); and West, Richard Williams (UCLA).

In order to provide opportunities for further review before conducting in-depth interviews in some LEAs, we adopted a two-stage interview strategy. In the first stage, five LEA interviews were conducted in each region (for a total of 20 interviews). In each case, one of the regional researchers contacted the district using the TAC and SEA referral as the point of introduction for the interview. These

initial discussions correspond fairly closely to the question agenda previously presented in connection with the SEA interviews.

Based upon the initial information about LEA evaluation utilization gained in these five LEA interviews, the regional interviewer and the project director selected two LEAs for further in-depth phone interviewing and one LEA for a district site visit in phase two. These 12 LEAs (three in each region) were, in our best judgment, the locations where the greatest utilization of Title I information was to be found.

The subsequent in-depth interviews lasted approximately an hour each. Respondents were selected from seven categories: (1) Title I Coordinator; (2) Title I Evaluator; (3) Superintendent or Assistant Superintendent (curriculum, instruction, program, etc.); (4) School Board Member; (5) Principal at one selected school; (6) Title I Coordinator or other individual primarily responsible for the Title I program at the same school; and (7) parent member of the school advisory council.

In some LEAs, certain categories did not exist, and those particular interviews could not take place (e.g., one person served both as district Title I Evaluator and Title I Coordinator). Conversely, in some LEAs certain of our categories of responsibility were shared by two people, and two interviews were conducted (e.g., there was both an in-house district evaluator and an external evaluator).

In each instance the regional interviewer phoned the individual he/she had interviewed in phase one and explained our interest in the extent of exemplary utilization within <u>that</u> district. The interviewer elicited the assistance of this LEA contact person to schedule phone

interview with the seven designated individuals. The process of arranging and conducting interviews was accomplished in a satisfactory manner. The researchers summarized the instances of utilization that were described on a district summary form that had been prepared for this purpose (see Appendix A).

In addition to the two LEAs selected for in-depth phone interviews, one district in each region was selected for a site visit. The selection was made jointly be the regional interviewer and the principal investigator. One full day and a portion of a second day were scheduled for each LEA site visit. During the site visits we interviewed the same categories of individuals that were selected for the phone interviews. However, the researcher on site had the opportunity to follow up on leads and interview people who had not been scheduled but who appeared to have important information based on the comments of other interviewees.

While the topics covered in the in-depth phone interviews and district site visits corresponded to the SEA interview agenda described above, the conversation went into much greater detail.

The regional researchers summarized the information on utilization in the four site visit LEAs on the district summary forms. More importantly, each researcher composed a case study narrative based on their face to face contacts describing the district, the content of the interviews and the insights they gained into Title I evaluation utilization. Particular attention was paid to the contextual factors that seemed relevant to the high level of utilization that occurred in these four LEAs.

CHAPTER 2

REVIEW OF RESEARCH LITERATURE AND TECHNICAL DOCUMENTATION

Introduction

This review of research literature and technical documentation examines the utility of Title I evaluation information for improving educational services at the local and state level, and it investigates the role of TAC support services in promoting increased use of TIERS data for program improvement. Among the many studies of evaluation utilization in the past few years, only six or seven examined Title I evaluation specifically. This group of research reports forms the core of our analysis. The studies we considered were all published between 1977 and 1981, and varied in scope from studies of a single school district to research on a national scale. Our data base included documents developed by staff members from ED, various TACs, SEAs, LEAs, and university researchers.[1]

Unfortunately, this body of research has one practical drawback as a basis for assessing the current impact of TIERS and TAC on local utility. Most of the TAC direct emphasis on local use of Title I data for program improvement has been too recent for its impact to be reported in the literature. Before mid-1980, TAC concern had focused most heavily on indirect increase in use presumed to occur if the

[1] We are grateful for the assistance of the TAC Clearinghouse located at the American Institutes for Research--TAC VII--in Palo Alto, California, for providing many of the documents on which this review is based.

quality of data were improved. Thus, efforts were directed towards refining procedures to improve the quality of Title I data. This was reflected in the construction of the TIERS system and in related training and technical assistance. It is only recently that direct efforts have focused on assisting LEAs to improve the use of evaluation information for decision making. The research studies that examine Title I data use are based on information collected prior to these recent efforts. Consequently, some indications of impact from TAC support services on LEA and SEA use of Title I evaluation data may only be derived from the field visits performed in conjunction with this review.

In addition to these various research studies about evaluation use, we have also reviewed a number of documents describing the evolution of TAC assistance over the same period. This broader understanding of TAC activities will help us assess the potential for increased evaluation use that can be derived from current TAC efforts, and will enable us to examine the relationship between past TAC activities and evaluation use in a broader perspective. There are many necessary conditions that must be met before evaluation data can serve the purposes of program improvement, and much of the effort of the TACs over the past five years has been directed toward establishing some of the precursors of use.

Consequently, this part of the report is divided into two sections. We begin with a review of the research literature on local use of Title I evaluation. We consider each study in turn, describing the major findings, and as appropriate, commenting upon the strengths and weaknesses of the analysis. Then we will review documentation related to

TAC support services. Specifically, we will present a discussion of how TAC activities might enhance the local use of evaluation data. This discussion sets the stage for the presentation of the state and local use study reports which will follow.

Review of Research Literature on Title I Evaluation Use

In 1978-79, the National Center for Educational Studies (NCES) conducted a fast response survey to determine school districts' attitudes toward the newly mandated Title I evaluation models (National Center for Educational Statistics, 1979). Sixty-two percent of the districts expected to need technical assistance in at least one new evaluation activity. Districts were concerned about a number of evaluation activities. The area in which they expected the greatest need for technical assistance was "evaluation for continuing program improvement."

This is a good place to begin our discussion of the research of Title I evaluation use, because the first studies suggested that the system was failing to meet this need and that TIERS data was not being used for program improvement. Only the most recent studies offer any evidence to contradict this pessimistic appraisal. We will discuss the research in chronological order, noting the extent to which our understanding has changed over time and how this conclusion requires modification.

Syracuse Research Corporation (1977) and Follow-up Study (1979).

The Syracuse Research Corporation (Goettel et al., 1977) study was conducted in 1976-77, prior to full implementation of the models and the

TACs. It provides a useful starting point for discussion of evaluation use in Title I. In addition, an updated paper reflecting recent trends was prepared in 1979 (Forgione, Kaplan, Orland, 1979) which provides a measure of how much change occurred after the implementation of TIERS and the TACs.

Forgione et al. summarized the conclusions about evaluation use of the SRC study. "The most important (of the study's major findings) is that evaluations were conducted for reporting rather than utilization purposes." In the eyes of their respondents, "state and local program evaluation was viewed most frequently as a mechanical exercise unrelated to other administrative activities."

While this negative appraisal held in most cases, there were some districts in which the level of evaluation use was higher. They examined three districts with marked evidence of evaluation use, and they found certain important commonalities.

1) The district's evaluation staff worked closely with the agency's Title I administrators and program officers.

2) There was considerable interest, if not actual pressure, emanating from the district's administration for more than a routine Title I program evaluation.

3) The evaluation thrust is not restricted to Title I programs.

4) There were experienced and talented evaluation staff people available in the district.

These studies preceded the implementation of TIERS and the TACs, and therefore serve primarily as background information, a baseline against which the efforts of the last four years can be measured.

In 1979, Forgione et al. tried to update their conclusions and to assess the impact of TIERS and the TACs by talking with district program staff with whom they had worked before. Based on these conversations, the authors concluded that "the most basic findings previously reported from these studies continue to hold true. Neither the models themselves nor the TACs have changed basic grantee attitudes towards evaluation utilization." The authors explain this partially by noting that the TIERS requirements had very little to do with utilization, and were primarily directed towards the provision of data.

On the other hand, they speculated that "the creation of the models and TACs have put forces into motion which may lead to increased utilization of evaluation information for state and local decision making." This hypothesis was based on the heightened awareness to evaluation that they perceived among school districts. "State and LEA administrators are being sensitized to the fact that evaluation is a serious and complex responsibility." They believed this awareness to be a necessary precursor to significant change. And they concluded on a positive note,

> TACs have themselves become a direct force for change by pushing for a larger role in helping grantees use evaluative data for their own purposes... It seems reasonable to conclude that circumstances now appear ripe for beginning to alter the persistent conditions observed in the SRC case studies concerning grantee utilization of Title I evaluation data.

SRI International (1978).

David (1978) conducted case studies in 15 Title I districts selected to have an above average emphasis on, or concern with evaluation. She sought to determine whether local staff used TIERS data to identify

program strengths and weaknesses in order to embark on program improvement, and whether the proposed changes in the Title I evaluation system then being considered were likely to alter the local use of evaluation.

David concluded, much as had Goettel et al., that "the primary function the evaluations served is to meet state and federal reporting requirements." Her respondents reported that "the evaluation ... results do not primarily serve as a means of judging the program or as a guide to program improvement."

David's analysis of why the evaluation results were not useful for program improvement was interesting, because it suggested that the upcoming changes in the Title I reporting system (selecting one of three mandated evaluation models) would have little effect on local utility of the information--as she put it, "...changing the type or quality of information contained in Title I evaluations will not, by itself, significantly affect local usage of these evaluations."

David analyzed how local school personnel made program judgments and decisions about program changes and found that other sources of information, such as skills related tests and personal judgment or observation, almost always carry more weight than standardized test results. She also noted that Title I programs are quite stable, and the changes that do occur tend to be marginal. "Thus, the universe in which to find connections between program change and evaluation is limited." In addition, she noted that evaluations were not timely, and that other factors such as the availability of funds and political considerations tend to play a larger role in decision making than evaluation.

Finally, she observed a lack of connection in many districts between the program staff and evaluation staff and found "little communication and understanding between those responsible for the administration and content of the program, on the one hand, and those responsible for the conduct of the evaluation on the other." This was one of the conclusions reached by Goettel et al. in the Syracuse study, as well.

In addition to these contextual constraints she also noted two additional constraints that limited evaluation use. She found that school staff perceived evaluation in a very narrow and potentially threatening way, "a set of procedures to provide one's superiors with information on which to judge the program, on the basis of criteria defined by those superiors." In addition, she noted that most Title I staff have a strong commitment toward the program and seek out evidence to reinforce their own positive assessments of the project, "effectively ignore(ing) evidence that does not support these feelings."

In looking toward the future and considering ways to improve the local use of Title I evaluation, David made three recommendations. First, "any strategy designed to increase local use of Title I evaluation must be grounded in a federal commitment to the school." There must be a widely disseminated and generally accepted understanding that government will continue to provide assistance and pursue efforts to increase local use of the data. Secondly, echoing Goettel et al. (1977), she suggested that assistance should be focused in the areas of increasing communication and cooperation between program staff and evaluation staff within districts. Thirdly, the Title I staff and the parents need assistance in "incorporating evaluation information into

planning and decision making. In particular, they need assistance in how to ask their own evaluation questions." David's point was that as long as the TIERS system seemed to be responding to external demands rather than local questions, the results would be perceived as "potentially more threatening than helpful."

David suggested the following strategy which presages much of what took place during the next year and a half in the area of TAC assistance to LEAs.

> The USOE evaluation models are designed to improve the quality of the data and will not, by themselves, lead to an increase in the local use of evaluation. However, the current technical assistance strategy, if redirected, could serve as a powerful force in changing how evaluation is perceived and thereby increase evaluation use locally. To accomplish this goal, technical assistance must be redesigned to communicate a new view of the role of evaluation and to develop skills such as generating one's own evaluation questions. As long as technical assistance is defined narrowly as a way of telling local staff how to improve the quality of their data, it will not increase local use of evaluations. (David, 1978, pp. viii, ix.)

Before we examine the next research study, we should consider the underlying definition of use that David employed in her judgments. Although David did not explicitly define what she meant by "use of evaluation" in her study, the questions that were asked in the interviews partially reveal her intention.

Respondents were asked, "How is the Title I evaluation used in your district?" They typically responded that it was used to meet requirements or to provide feedback to school staff and parents or to provide a rough index of the program's impact on achievement. This is an "instrumental" notion of use--having an identifiable impact on a specific action. Such questions probably would not elicit indications of incremental, gradual influence on attitudes and judgments, and these, too, are important components of utilization.

On the other hand, in subsequent questions respondents were asked about how they formulated judgments about program effectiveness, and about how they would go about deciding on program change. Such questions are more sensitive to a broader notion of evaluation utilization. Yet, it is not clear from the report how wide a conception of use underlay the questioning and analysis. What was described in this study and in the Syracuse Research Corporation study as well, leads one to suspect that many aspects of evaluation use in its broadest sense were overlooked or ignored--tending to underplay the actual impact of Title I evaluation.

Hamilton/Chicago (1979).

Hamilton's (1980) study of evaluation in a single urban school district had findings similar to those already discussed. She noted that, "one of the most striking findings was that few administrators see the purpose of evaluation as improving Title I programs or policies. The most frequently mentioned role of evaluation was that of an auditor."

This study was conducted on a school district that had recently changed the emphasis of their evaluation and research efforts. During the year of the study, they shifted focus from national information requirements to evaluation for individual school oriented decisions. While some changes in attitudes and practice could be seen as a result of this shift in emphasis, Hamilton still found that, "Title I evaluation input rarely affects Title I policy on the local school level." Hamilton herself noted the similarity between these findings and those of David (1978).

At the same time, Hamilton noted that the local focus of some of the district's evaluation efforts had begun to have a positive impact on the attitudes expressed by school principals. "Most administrators appreciated evaluation efforts when they took the form of discussion and planning between the evaluator and the principal... The evaluation design change which the principals most frequently indicated would be helpful on the local level was to increase even more the opportunities for feedback and planning between the evaluation department and the school." These comments repeat the suggestions of Forgione et al. (1979) and David (1978) of the importance of bridging the gap between program personnel and evaluation personnel at the local level.

Hamilton's conclusions reinforced the picture of evaluation use that emerged from the previous two studies. However, Hamilton had an even narrower conception of "evaluation use" than David or Goettel et al. Hamilton's research was concerned with local decision making, and decision making included only the selection of new Title I programs. Thus, Hamilton did not consider formative program changes nor long-term questions such as the formulation of attitudes that might have impact outside the immediate choice of Title I programs.

Northwestern University (1980).

In a comprehensive study of evaluation practices and procedures in federally supported programs at the national, state and local levels, Boruch and Cordray (1980) devoted some attention to the local use of Title I evaluation data. Their conclusions were drawn primarily from other published sources, and they found nothing to contradict the impression we have conveyed so far. In fact, they add some additional

weight to these conclusions based on the site visits they conducted. "Our own site visits also suggested test scores are not and should not be a basis for major local decisions about a program. Most decisions involve small modifications in any case. The information used in small modifications includes test scores, incidental observations, an attitude survey, teacher or parent criticism and others." This statement reminds us of David's observations that only marginal changes occur in most Title I programs, and that other information is frequently more salient than test scores.

Boruch and Cordray brought up another issue of great importance to this study--the degree to which LEAs supplement the required Title I evaluation with additional evaluation activities of their own. Excluding small districts, they found that at least three quarters of the LEAs' evaluations "go beyond federal reporting requirements to some degree." The most typical supplemental evaluation activities that were reported included: "one time only surveys of teacher attitude, assessment of changes in the affective domain produced by exposure to inservice workshops or the Title I program, classroom observations, and assessment of locally developed material." More sophisticated districts, according to Boruch and Cordray, performed some studies assessing validity of achievement data, some evaluating implementation, and some determining whether one program environment was different from another. Since such supplemental studies were not the main focus of their inquiry, they did not have an extensive review of the impact of this type of evaluation. However, their comments are important because many have advocated that utilization can be enhanced by encouraging

districts to supplement the minimum TIERS requirements (Alkin, 1980). According to Boruch and Cordray, some of this supplemental evaluation activity was already being conducted in most school districts. "For a substantial number of districts with enrollments of 5,000 or more, there appears to be at least some effort devoted to obtaining information that is suited to the needs of the district personnel."

The Huron Institute (1980).

The Huron Institute was commissioned by the U.S. Office of Education and the National Institute of Education to conduct a two year study on the role of evaluation and testing information in public schools (Kennedy, Apling, and Neumann, 1980). One particular aspect of that study was to examine the role of evaluation and testing in Title I programs. The report from the first year of that study (Kennedy, Neumann, and Apling, 1980) shed some interesting light on the question of the use of TIERS data and TAC assistance. The major conclusions that were reached are not markedly dissimilar from those already reported, yet this more recent study approaches the questions more broadly and in greater detail than any of the previous efforts.

Kennedy and her colleagues conducted first-hand interviews in 18 school districts identified as districts that did use evaluation or test information in decision making. The respondents were asked to discuss the issues that most concerned them in their current jobs and to describe what information, if any, they relied upon in making decisions. This is a broader approach than was taken in any of the earlier studies, and is potentially more sensitive to subtle aspects of information use.

In their analysis of the interview data, they differentiated between three broad sets of individuals: those who focused primarily on district policy, those with general program responsibilities for Title I, and school building personnel such as principal and teachers. The three groups used information somewhat differently, but some overall conclusions applied to all three groups. The most important of these was that evaluation use and the use of test data were organized around the type of issues faced by the school personnel. The local concerns of the personnel dictated their interest in and attention to evaluation data. Boruch and Cordray (1980) had previously made the point that the needs of the local personnel were important determinants of use, and the Huron study confirms this notion.

District level policy makers had very little use for Title I data because most were not concerned with issues to which such data were directly relevant. The issues of most concern to them were federal policies, changing enrollments, student behavior, and testing. As Kennedy, Neumann, and Apling noted, "the information that was most often relevant at the district policy level consisted of descriptive data bases." In fact, the, authors added that, "although Title I funds are often the largest single categorical grant a district receives, none of our policy making interviewees described Title I as a major policy concern."

Similarly, and possibly more important, Kennedy, Neumann, and Apling found that Title I program directors were most concerned with their own responsibility and authority, the overall level of funding of the program, and compliance with federal, state and local program

regulations. These are not areas in which TIERS data are relevant, and it should not come as a great surprise that such data are not used by Title I program directors.

However, the Huron study went further in its examination and found these same program directors did use locally designed evaluation results. The investigators found (as had Boruch and Cordray) that both formal and informal locally designed evaluations were often a source of useful information. "Although the studies that were used were not mandated by or shared with the state agencies, they were nonetheless supported by the district's Title I grants."

Such evaluation studies had three different purposes. The first purpose was the director's need for knowledge of the program. Typically, formative evaluations of the Title I project met this need. The second information purpose was compliance, and often resulted in informal or process evaluations of the number of students served and the types of services provided. The third and most interesting purpose was assessment of program delivery strategies; this prompted studies in which a district compared two alternative strategies for conducting certain activities. There were not as many of this latter type; Kennedy and her colleagues noted, "While we found fewer examples of studies comparing alternative service delivery systems than we found assessing compliance, the fact that any were done is encouraging."

The Huron study looked further to determine the elements that influenced the use of information by program managers. They found that the manager's personal knowledge about the program and his or her

relationship with the evaluator were more important than any specific program results. The first of these observations jibes with similar conclusions from the SRI International study and the Northwestern study that personal information may be more important than test score data. The second supports the observation made by David (1978), Forgione et al. (1979) and Hamilton (1980) that communication between the program personnel and the evaluation personnel is a critical determinant of evaluation use.

Strong evidence for the use of Title I test data came from teachers. Teachers used tests of different kinds in a number of ways. Tests played an important role in selecting students to receive Title I services, in identifying the particular educational needs of the students, in evaluating their overall progress, and in planning program improvement. In some cases, it was the standardized TIERS test data that were used; in other instances, teachers used supplemental evaluation instruments which were not part of the mandated testing for Title I. The use of supplemental Title I evaluation procedures, such as criterion referenced tests and other local instruments, were particularly relevant for identifying student needs.

Since there was a marked difference in the use of test data between teachers and program managers, it should not be a surprise that the concerns of Title I teachers were also considerably different from those of the program managers. Teachers were worried about student selection, testing out of the program, and students' needs. Teacher use of the data was influenced by the accessibility of the data, their principal's interest in test data, their familiarity with test data, and their

professional commitment to data use. We note in passing that at least two of these four factors might be influenced by staff development and technical assistance.

The Huron investigators also spoke with evaluators, and determined the issues they identified as important. The evaluators described a recent change in several districts to more centralized evaluation services. "We found evidence of this trend toward separating program evaluators and program directors in several of the districts we visited." Many evaluators interviewed in this study noted that "the nature of the relationship between a program director and the program evaluator was a critical factor affecting the extent to which evaluative information would be used constructively." They characterized the conflict quite succinctly as, "greater objectivity through separation versus increased understanding through close involvement."

The Huron study also offered some specific conclusions about the influence of federal policies on the use of Title I information. First, they found that Title I staff are relatively sophisticated in their understanding of the values, uses and limitations of tests. The authors made the point that both Title I program directors and Title I teachers had a more sophisticated understanding of evaluation and testing than other program directors and teachers, "a finding we attributed to their years of experience with these data." This is precisely the kind of heightened awareness that Forgione et al. (1979) felt was a necessary condition for improving evaluation use.

Second, they noted that "while the regulations do not specify evaluation beyond needs assessment, the models, and the tri-annual studies

of sustained gains, they encourage other studies and nearly every district had invested some of its Title I funds into areas other than those explicitly required."

Third, they observed that the confusion over the evaluation models that had dominated much of their attention in the first two or three years of their use had subsided considerably. In fact, normal curve equivalent (NCE) scores had even found their way into evaluations of non-Title I programs.

To summarize, the Huron study found little evidence of direct use of the standardized test data in the 18 sites they studied. However, they did find a number of instances of use of other Title I funded evaluation information. This is encouraging. However, it is important to note that this finding does not represent the state of the art nationally. Rather, since the study focused on identified "exemplary" districts, the findings must be viewed as an indication of utility--the possibility of use.

Finally, it should be noted that impacts examined in the Huron study added a caveat that there are more incremental notions of use, and other kinds of evaluation influence that were not examined in their study.

State of Indiana (1980).

In 1980, the State of Indiana was awarded a refinement contract to study the use of TIERS data for decision making and program planning in LEAs across the state. This study included: a questionnaire to examine evaluation practices statewide; a set of audits to determine the accuracy of TIERS data reported by LEAs; and a set of ten case studies of Title I evaluation practices (Crowley, 1980).

Based upon the questionnaire, Crowley concluded that there was strong evidence of evaluation use: 92% of the LEAs that responded reported that they used TIERS information. However, there was no specific description of how the information was used, nor what was meant by the word "use." Subtler understanding of the "uses" in Indiana LEAs had to be derived from the case studies.

In four of the ten case studies of LEAs, TIERS "performed a major role and provided critical information for major instructional program changes." This is an extremely powerful statement, stronger than in any other study we reviewed. Unfortunately, it appears to us that the author is unjustifiably emphatic. In two of the four cases "TIERS supplied objective information which confirmed the decision that change was needed," while in the other two instances "TIERS stimulated local decision makers to examine the Title I program closely and determine a need for a change." In these four instances, the use of the information was in recognizing the need for change, or confirming a nascent recognition that there was a need for change.

Lest we be too critical, Crowley observed several other uses of TIERS. Seven of the ten Title I coordinators "viewed TIERS as a measure of overall effectiveness, or a global indicator that something is right or something is wrong."

The Indiana study also showed that TIERS influenced decision making as one of multiple evaluation inputs. (Both the Northwestern study and the Huron study noted similar use of other data to supplement TIERS.) Nine of the ten LEAs indicated that TIERS data helped them to determine program strengths and weaknesses. However, in

every case, "multiple sources of information were used to make this determination." In addition to TIERS data, districts also used teacher feedback, informal evaluation, brief surveys, parent input, etc. They cited a number of indicators of program strengths and weaknesses in addition to test scores. These included improvement in student attitude, information gathered from personal expertise, students' interest and willingness to attend Title I projects, and an increase in students' self-concept.

While Kennedy, Apling, and Neumann (1980) reported a strong use of TIERS data by teachers, Crowley confirmed other sources of information that teachers used in decision making.

> The following items were consistently mentioned by teachers: a) improved student behavior related to the teacher and other students, b) achievement impact evaluation, c) local mastery tests and textbook mastery tests, d) increase in self-initiated and independent work on a student's part, e) increase in positive attitude towards school.

It is interesting to note what factors the Indiana LEAs suggested were hindrances to their use of TIERS evaluative data. The first was that they were unfamiliar with NCEs and did not have a general context for interpreting them. The second dealt with the efficiency of implementing TIERS. Third, was that the LEA personnel did not have sufficient knowledge in tests and measurement and had no confidence in the results of TIERS. Similar factors were observed in the Huron study. We mention these because they are all areas that can be directly affected by technical assistance currently underway by TACs.

Crowley concluded that TIERS information is used to confirm or stimuate instructional program change, and that LEAs attempt to systematically determine the program strengths and weaknesses utilizing multiple sources of information. He added that with further experience and knowledge, "more confidence will be gained, use of TIERS will increase, and the quality of use will improve" This echoes the sentiments of Forgione et al. (1979), and Kennedy, Apling, and Neumann (1981) that heightened awareness and familiarity will lead to increased use.

RMC Corporation (1981).

Cichon and Herr (1981) performed an in-depth examination of the use of information from a supplemental Title I process evaluation conducted in a northeastern urban school district. This study has limited value for our purposes because it was not an investigation of TIERS data or of the normal TAC assistance, but rather a specially funded, system-wide process evaluation. On the other hand, it is important because it indicates the potential for evaluation use that can be achieved under the right circumstances.

One distinctive feature of the process evaluation under study was that it was "greater in scope than the Title I evaluation and reporting system requirements." In addition, "the program staff was involved in defining the issues of the study." Thus, such an evaluation is beyond the range of what occurs in the typical school district. Nevertheless, it exemplifies how much utilization can be enhanced by a cooperative effort between program personnel and evaluation staff.

Cichon and Herr found "evidence of high utilization" of the information from the process evaluation. They hypothesized a utilization continuum running from incremental at one extreme to synoptic at the other. Slightly over half the uses they reported were the incremental type. This suggests that previous research may have overlooked a great deal by ignoring gradual impact of TIERS data. At the same time that they were noting these gradual uses of evaluation, they also noted that 53% of the specific recommendations that had been made in the process evaluation were implemented. Thus, there was also very direct evidence of use.

Center for the Study of Evaluation (1981).

Stecher, Alkin, and Flesher (1981) reported the results of an interview study of information use among 65 elementary school principals, Title I coordinators, and resource teachers from a single urban school district. All schools involved in this study were receiving Title I funds, though many had other special supplemental programs as well. Stecher and his colleagues allowed the respondents to identify significant occurrences in the life of the school's special programs and used these events as the basis for further questioning. They examined decisions associated with the identified occurrences to determine the role played by different types of information. Their findings corroborated a number of the observations made in previous studies.

First, they found that formal evaluation data were used only infrequently in decision making. Personal opinions and beliefs of the participants were the most salient factors in the decisions. Other often

cited inputs were program regulation and guidelines and informal evaluation data. These results substantiate similar observations made by David (1978), Boruch and Cordray (1980), and Kennedy et al. (1981).

Second, they divided the decision making process into four phases--problem recognition, decision interaction, ratification, and dissemination--and found that information use patterns differed between the phases. Specifically, they found greater reference to evaluation in the recognition phase. This finding corroborates the observations on the use of Title I evaluation made by Crowley (1980) in the Indiana case studies. By making a distinction between different stages of the decision process, the CSE study was able to identify the important role that evaluation plays in problem recognition, and differentiate this from the minor role it seems to play in the selection of the final course of action.

Third, they found a relationship between the type of information that was used in a decision and the type of issue under consideration. For example, there was little or no use of evaluation in administrative decisions, while there was more frequent (though still limited) use of evaluations in decisions involving broad curriculum issues. Kennedy et al. (1980) noted that the issues of concern to the individuals affected their use of information, and this result confirms the relationship between this type of decision and use of information.

The latter two results of the Stecher, Alkin, and Flesher study have an additional important implication for this project. They suggest most strongly that it may be impossible to find broad generalizations about Title I evaluation use, and indicate that any study of this area should be sensitive to the different types of decisions that are being considered and the different stages that exist in most decision processes.

Summary

The pattern of results in the various studies of the use of Title I evaluation information is clear and quite consistent. There is a widespread confirmation of the finding that Title I evaluation is primarily directed toward reporting and compliance functions. However, early appraisals that this was the only use of Title I evaluation have been moderated. Current research offers a more expanded point of view. Though TIERS remains in a compliance oriented system, literature provides evidence that local use is being made of other Title I funded evaluation data. These data derive primarily from supplemental local evaluation activities funded by Title I. The research indicates that various types of supplemental Title I evaluation help local school personnel recognize areas that are in need of attention and confirm impressions that have been obtained through other sources. In addition, there is evidence of other indirect and incremental uses of Title I evaluation (of all types) which are much more difficult to characterize (Cichon and Herr, 1981; Stecher et al., 1981). Overall, these studies indicate that there is considerable use of supplemental Title I evaluation information for problem recognition, though there is little evidence of direct use of evaluation data per se.

Thus, the research does substantiate instances of utilization of Title I evaluation data. However, most of the documented utilization is of the indirect or incremental type, and there is little evidence of TIERS impact.

The Role of TIERS and TACs in Local Evaluation Use

While the major research studies of the last four or five years present very modest evidence of TIERS evaluative data use for local program decisions, there are many reasons to believe that the situation is changing. Conclusions reached in prior research studies emphasized that the implementation of the TIERS models themselves had limited potential for evaluation use, and it was only through subsequent efforts on the part of the TACs that any improvement in evaluation use could be anticipated.

During the last school year, there has clearly been a major emphasis on local use in the efforts of the TACs to assits LEAs and SEAs. As Stonehill and Groves (1982) noted, TAC activities had changed dramatically in the last five years. "Early TAC efforts seemed to focus on assisting SEAs and LEAs to understand the guidelines and procedures behind TIERS." This made very good sense because the new system was unfamiliar to LEA and SEA personnel and their first and foremost need after the models had been introduced was for help to understand the procedures, and the proper way to handle data collection and reporting.

Stonehill and Groves characterized 1976 as a "year of orientation," and 1977 as a "year of the technical issue" (during which time there was an extensive national debate over the theoretical and statistical assumptions underlying the TIERS models). By the same token, a focus on quality control issues characterized 1978. This concern with the quality of the data that were being reported grew out of the realization that there were many mistakes being made in scoring data trans-

formations and analyses. This focus on quality control "included the development of computer programs that scored and screened data to improve their validity, the development of hand calculator software, the introduction of optical scan sheet scoring systems, etc."

Stonehill and Groves then characterized 1979 as the "year of local utility." In fact, these efforts, which were begun in ED in 1979, were just beginning to reach the SEA and LEA levels during the 1980 school year. A review of recently developed TAC materials suggests that emphasis on the quality of data prevails as a major TAC concern, but that concerted efforts to improve the local use of TIERS data have begun to emerge. Thus, full evidence of the result of that effort will not be found in published research studies.

By indicating that a concerted focus on evaluation use is a recent phenomenon among the TACs, we do not mean to suggest that their earlier activities were irrelevant to this goal. In fact, much of what the TACs have accomplished over the last four years may well have been necessary precursors to any greater local use of evaluation from Title I. Taking a long term view, Stonehill and Groves pointed out that increased evaluation utilization is clearly a long term objective. After the new system had been introduced, it was necessary for the TACs to help LEAs and SEAs implement it correctly, learn to use the requirements, improve the quality of their data and eliminate errors in scoring, etcetera, before it made any sense at all to try to encourage the use of the information for decision making. Each of these was in fact a necessary step to take before the district could hope to make use of the information for program related decisions. The old dictum that

you can't have valid measurement unless it is reliable applies to TIERS as well; reasonably high quality data has to be produced in order to have any credibility in the eyes of local audiences.

Similarly, Millman et al. (1979) concluded that "many SEAs and LEAs would like more emphasis placed on quality control and on making evaluations useful for the local level. There is agreement on the continuing need to upgrade the capability of SEAs and LEAs in program evaluation so that evaluations are conceptualized and conducted in a manner that provides useful results at those levels."

A similar discussion will be found in Anderson, Fishbein, and Stonehill (1979). They described the two-pronged effort of the Evaluation Use Committee in ED to focus both on improving product evaluation and on developing models for process evaluation. "The goals of this project are to develop an orientation towards asking meaningful questions about programs, to emphasize the need for quality data, to show how evaluation data generated from TIERS can appropriately be interpreted, to illustrate the planning and implementation of a formulated evaluation strategy, and to provide actual examples of how evaluation data have been used in decision making." Materials related to these two strategies have been produced by various TACs, and workshops based on these materials were first conducted during the current school year. We note positively the correspondence between this approach and the recommendations of the studies summarized in the first section.

As far as can be determined, TACs have been extremely successful in their efforts to date. In a TAC initiated study, Hansen and Oxford (1979) found "TIERS and the TAC system have been instrumental in

improving Title I evaluation across a broad front." They note improvements in test selection procedures, test interpretation and scoring, forms and reporting procedures, and general evaluation practices. More importantly, harkening back to the considerations of David (1978) and Kennedy, Neumann, and Apling (1980), they note a dramatic improvement in attitude towards Title I evaluation. They attribute this to the positive personal contact that LEA and SEA personnel are having with representatives of the TACs. "While the TIERS system can provide a mechanism for data collection, aggregation and reporting, it cannot by itself positively influence attitudes towards Title I evaluation... These improvements in attitudes then can most likely be attributed to the human element in the TAC delivery system."

Hansen and Oxford continue quite reasonably to ask what then can be done beyond improving attitudes and quality of data to influence the actual use of evaluation. "The results of the study indicate that significant progress has been made in improving Title I evaluation; however, there is much more to be done than has been accomplished to date. Training programs in evaluation planning and in the use of evaluation data in decision making and supporting materials must receive a greater emphasis than they have in the past."

Troy (1979) describes this effort in somewhat more detail. Four activities followed in the wake of the formation of the Evaluation Use Committee. The first was the development of an annotated bibliography on utilization; then a detailed review of literature was completed. A set of case studies related to evaluation utilization was compiled, and finally a series of workshops were developed to help LEAs improve utilization

skills. The workshops grew out of the case studies, and represented a specific set of TAC activities designed to go beyond merely improving the quality of data to translating this good data into local use.

Of course, TIERS was never intended to be the only source of information for local school decision making. Stonehill and Groves (1982) note that "the incorrect inference was the USOE intended that evaluation should consist exclusively of aggregating results from tests." In fact, local LEAs and SEAs are supposed to conduct whatever additional evaluation activities they deem necessary to help them make decisions about programs.

In seeking to encourage LEAs to increase evaluation use, TACs have extended their technical assistance activities to emphasize other types of evaluation beyond the mandated models. The extensive process evaluation reported in Cichon and Herr (1981) for example, was a TAC supported activity. More important, it indicates a trend in TACs to assist LEAs to identify local problems and develop evaluative strategies to address them.

The conclusion to be drawn from this review is that the TACs have undertaken a number of activities in the last year that show promise for improving the use of Title I evaluation for program improvement. The reports produced by ED and the TACs provide ample evidence of necessary precursors to evaluation use that have been addressed by the TACs in prior years. More importantly, our review of the research literature indicates that the current efforts of the TACs seem to be proceeding in the right direction. Current TAC activities--the workshops that have been developed and the materials that have been

produced--appear to be in line with recommendations in the literature on Title I evaluation use, and one would anticipate that continued efforts in this direction could have a positive impact on the use of Title I evaluation for program improvement.

CHAPTER 3

TITLE I EVALUATION ACTIVITIES AND USE--SEA LEVEL

SEA Title I Evaluation

This section details our observations based upon interviews with State Title I coordinators and evaluators. As noted above, we selected 12 SEAs for interview--three each from the east, midwest, south, and west. The purpose of these interviews was to document generally the evalaution functions performed at the SEA level, but more importantly, to attempt to determine indications of evaluation use.

We will first provide a brief description of the primary evaluation activities conducted at the SEA level, and then discuss the organizational location of the evaluation.

Evaluation Functions

We found three major kinds of evaluation activities at the SEA level. Working with an expansive view of evaluation we found that evaluation is not simply restricted to the collection and analysis of TIERS data. Functions also included various monitoring activities conducted by the Title I staff, and technical assistance to LEAs.

TIERS. The examination of the SEA interviews revealed, obviously, that in all cases TIERS data are being collected and that this is viewed as the primary responsibility of the SEA evaluation personnel.

However, in approximately one-third of the instances, additional data not required in the Title I Evaluation and Reporting System are specified by the state agency as important, and included as a reporting requirement for local districts. This kind of data typically add more detailed cost data than required by the federal regulations. For example, one state of those surveyed requires cost data broken down by sub-categories, such as parent involvement, administration, instruction, and supplemental services. Other SEAs add certain indicators. For example, one state requires LEAs to collect data on the extent to which the school has a management plan, and an internal assessment system for managing student progress. Apparently, research conducted in that state has led state officials to believe that these two variables are critical for higher levels of educational outcome. Therefore, they seek additional evaluation information on these dimensions from each Title I project. Several states we surveyed include various items of non-required process data that they consider essential.

Several of the states collect data on an individual student basis rather than school or district summary data prepared by LEAs. Typically, in those states, information for each student indicates student's grade level, sex, race, and some description of the instructional treatment provided for the student. One midwestern state, for example, codes data for each student on the number of minutes in the program.

Most of the SEAs studied use a computer for aggregation of data at the state level as well as for the preparation of summary reports. In addition to preparing the required federal reports, every state also produces district summary reports. In all instances, these reports

summarize the outcome test data, but some summarize process data as well. Many states have prepared much more detailed reports with summaries by grade level in projects (typically a school building site), and by grade level within districts. Some states have prepared more focused reports with summaries provided in terms of categories of districts (size, economic characteristics, etc.). Several of the states in this study have produced separate reports specifically geared towards the urban districts.

Monitoring. The monitoring functions are typically performed by consultants within the Title I office--although there are a number of minor variations on this theme. In several instances, monitors have regional responsibilities. In other cases, the staffs required for monitoring are quite small and hold other responsibilities as well. Monitors visit schools primarily to insure program compliance; however, monitoring activities are also designed to provide program advice and consultation. Typically, the specific advice provided emanates from the expertise and experience of the monitors rather than from the evaluative data.

Technical Assistance. A variety of evaluation related technical assistance activities are provided by SEAs. One of the most prevalent types is designed to insure that the TIERS data are collected and provided properly by the LEAs. Other technical assistance is related to developing the District applications, particularly the evaluation-related portions of the application (such as the needs assessment data). Technical assistance is provided for compliance type issues such as the assurance that student selection at the LEA level is taking place in an appropriate manner.

We found one other type of technical assistance to evaluation. Some SEAs have engaged in locally-based development of procedures, instruments, and forms which might be of value for local information purposes. For example, one southern state interviewed during the course of this study has devised a set of 22 data collection charts which districts might use in conducting process evaluations.

Other Evaluation Activities. This study uncovered another interesting evaluation activity in an eastern state where LEAs are required to submit two additional evaluation reports. One of these, referred to as an "Implementation/ Interim Report," is particularly interesting given our concern for the use of evaluation information. The implementation report must discuss the start-up problems for the project, the data related to the extent of implementation, and the extent to which the LEAs have "attended to" the recommendations of the previous year. In essence, the report requires districts to indicate the use made of one portion of the previous year's evaluation findings--namely, the recommendations included with the district evaluation report.

Organizational Location of Evaluation Functions

The monitoring and technical assistance functions, are, in all instances, organizationally located within a program-oriented office. In some cases, this is a Title I program office; in others, the name of the office might be broader and include all federal programs. These evaluation related functions are primarily administrative and instructional. The administrative aspects have to do with the allocation of resources and monitoring to assure that federal regulations are being complied

with. The instructional functions deal with the provision of instructional program advice to LEAs--primarily in response to low outcome scores for the project.

When we turn to the TIERS data and its collection and analysis, we find a more varied picture. This function, which is more typically thought of as evaluation in the narrow sense, is handled in a variety of ways within the states that we examined. In four of the twelve states, this "evaluation" function is housed within a research and evaluation office and not within the Title I unit per se. In the eight other states, the evaluation function is in one way or another housed within the Title I (or federal) program unit. In at least four of the instances, one individual has the specific evaluation responsibilities. In several cases, this individual is an evaluator for the federal programs office, and in two instances the evaluator is specifically responsible for the Title I. In two other cases, the evaluation responsibilities are handled by specialists who have other duties (e.g., program monitoring). In one of these states, the evaluation responsibilities are shared among the specialists, while in one state the Title I responsibilities are handled by the State Title I Director. In two of the twelve states examined, the evaluation is performed by an external contractor hired by the state; in one of those instances, the Title I director took responsibility for delineating the evaluation and acting as primary liaison, while in another state a person with "full time evaluation" responsibilities within the department had the external liaison function.

Thus, one clearly notes that there are a variety of ways in which SEAs organized for the accomplishment of evaluation activities. There

are, of course, advantages and disadvantages to each of the arrangements, and these will be discussed in a later part of this section. One pattern is clear: in the large states the evaluation function tends to be centralized within a research and evaluation branch, whereas in the smaller states evaluation tends to be a much more broadly held responsibility and is administratively housed within the Title I program unit.

SEA Evaluation Use

We have distinguished three categories for state uses of evaluation information: 1) compliance, 2) problem recognition, and 3) problem solution. This section describes each of these types of state use in further detail.

Compliance

Each of the SEAs reviews and uses LEA procedures and reports to determine if they have complied with Title I requirements. Among the SEAs we surveyed, the accuracy of the achievement data receives the greatest attention. For example, one western state conducts a "Spring Swing" on the heels of the LEA spring testing to review data collection and conversion procedures in LEAs and to spot quality control errors. Similar direct action to improve the quality of the data was reported in most states.

However, in one case the SEA preferred an indirect approach. According to the state official responsible for Title I evaluation, "The mechanism for insuring accuracy of data was the promise of data back to districts. Thus districts in valuing that report would be inclined to send data which are accurate."[1]

Proper implementation of the models was another compliance issue that received a great deal of SEA attention. During the past four or five years, both SEAs and TACs have provided a great deal of assistance to help districts understand the models and use them correctly. As one southern SEA described it, "During the last five years or so, we have been in the orientation state. Inservice sessions have been held at least once a month at which we distributed materials (technical manuals, technical reports)." As a result of this continuing concern with proper use of the models, almost all SEAs examine district Title I reports carefully to determine if each LEA has correctly applied one of the three Title I models.

Other compliance issues reviewed by the SEAs include student selection criteria and Title I application procedures. For example, a western state has paid particular attention to the manner in which Title I participants were identified to be sure that LEAs were providing service to the "right students." Another state had regional officers use previous years' evaluation reports and needs assessments to help counties write up new applications correctly--yet another kind of compliance use of evaluation.

In several cases, compliance reviews are the main SEA use of Title I evaluation. The state agency limits its concern to compliance and passes to the LEAs the option of using the data for other purposes. One SEA summarized the situation in this manner: "If there is use (in

[1] This is an example of common differences between SEAs that we found throughout our review. SEAs differed in their willingness to intervene directly in LEA activities. Some adopted a rather passive stance, leaving it up to the individual LEA to decide what actions to take or what use to make of information. Others were active in encouraging or even mandating LEA behaviors.

this state), it is at the district and/or school level. At the SEA level it is pretty much just compliance."

However, this was not the most common situation reported. Most SEAs extend the scope of their review to include assessments of the quality of the programs that are being conducted by the LEAs. As we will describe in the next section, the Title I achievement data and other supplemental Title I evaluation data form the basis for judging program strengths and weaknesses.

Problem Recognition

In most cases, SEA compliance reviews are part of a larger program monitoring and review effort. In addition to data quality and compliance with the Title I regulations, such reviews also examine the overall success of the program.

States typically compile district by district summaries of outcome results and make these reports available to LEAs. They hope that individual LEAs would use these data to judge the strengths and weaknesses of their program and take action to improve areas that need attention. Thus, for example, a district would receive a report of its own achievement scores compared to the avergage scores statewide; the latter data would serve as a baseline for judging the success of the district's Title I program. An alternative strategy employed in some states is to send a district its own scores and the average scores of other districts with "similar" characteristics. In either case it is hoped that the LEA will make use of the comparative figures to identify problems or judge success of their own programs.

In some cases, SEAs view the passive role of sending out this information as their primary responsibility. They provide such comparative achievement scores, but do not require any particular action on the part of the LEAs. "The state takes responsibility for the evaluation reports, but then leaves it up to the LEAs," is how one western SEA described this outlook. Another SEA explained that they routinely send copies of evaluation reports back to each district in the hope that this data would be used by building principals. "The emphasis is on making school building principals into instructional managers. We believe that this comes from the availability of achievement information at the building level." In these cases the state agency downplayed its own role in bringing about changes in the LEAs' Title I programs beyond providing them with the data. In fact, several states require LEAs to conduct supplemental process evaluations for their own use.

However, in most cases the SEAs take a more active role in encouraging local districts to act on Title I evaluation data. The emphasis of these efforts is to identify the strengths and weaknesses of LEA Title I programs. Low scores, poor performance, etc. are used as indicators of problems with the Title I program. In this manner, the Title I evaluation is used extensively for problem recognition. Such a situation occurs, for example, in an eastern state where the districts with particularly low scores are "flagged" by the evaluator, and this information is shared with the four regional program offices that do Title I monitoring and review. SEA program officers can use this information in their personal interactions with districts. Similarly, many states make the Title I performance data available to their monitoring teams before

they visit schools. This information helps them identify potential problem areas and focus their attention during their limited time in the district. In one case, the state uses the reports, "to identify schools where there are larger clusters of low achieving students. They review this . . .data with the staff and attempt, in addition, to be responsive to other needs presented by that staff."

We should add that it is not just poor performance that brings districts "recognition" from the SEA. In one mid-eastern state, the SEA reviews Title I performance data "to try to pick out districts which, over time, have produced good results." These districts are encouraged to apply for JDRP, and the SEA assists the district in this process. Thus, the Title I data served not only for problem recognition, but also for identification of model programs.

Finally, we should note that while most SEAs use Title I data to assist in problem recognition (or identification more generally), some would prefer to have other indicators. One midwestern state was very direct in indicating its preference for its own objective referenced tests to the instruments used for Title I evaluation. TIERS was not well accepted because the state's director of evaluation felt that their own assessment program was more advanced.

Problem Solution

The most easily recognized use of Title I evaluation data occurs when the SEA incorporates the information into a problem solution. We noted above that some SEAs perceive their role as providing information back to the LEAs, while leaving the burden of action at the local level. Equally as often, however, the SEAs respond to the problems that were identified and take actions to correct them.

The most common use of Title I data in problem solution were of three types:

1) initiating programs of inservice and technical assistance to LEAs;
2) making programmatic recommendations to districts based on systematic study of factors that affected educational outcomes; and
3) using that data as input to some other educational question that was being decided at the state level.

Technical Assistance Uses. An example of the first type of use occurred in a western state where the evaluator sat in on planning sessions and brought knowledge of the aggregated Title I performance data to bear on relevant issues. In one typical case, this information was used to decide "areas for inservice training with respect to basic skills." A number of other states use Title I evaluation to locate areas of need for technical assistance. Sometimes they undertake the inservice themselves; other times they seek assistance from TACs.

However, SEAs typically reported that they based their technical assistance on more than Title I evaluation data. An eastern state indicated that they use the Title I evaluations for identifying districts with problems, but that the regional reviewers only make minor use of the data in deciding what advice to give to LEAs. "From the data, we identified the bottom 25 districts and sent teams out to help them." They reported, however, that the information gleaned from the state testing provides problem recognition only, adding, "Advice comes from on site observation."

Finally, one midwestern state cautioned that such warning signs are not always heeded by SEAs. They acknowledged that "poor achievement data is a red flag," but added that "it would almost have to be at the extreme to get us started."

Program Monitoring and Recommendations. There were many examples given of SEAs use of Title I information to analyze program characteristics and make recommendations for problem solutions. In some cases the SEA relies on the basic Title I performance data, but in many states the SEA or the state Superintendent of Instruction has imposed additional evaluation requirements on Title I LEAs, and this additional data is utilized in conjunction with the basic Title I achievement data for guiding state action. Some states base their recommendations on casual study of the data; others conduct systematic review and formal analysis before determining their preferred course of action. Some states offer these guidelines as informal suggestions to LEAs and allow each district to decide for itself whether it would adopt the particular recommendation; other states make their proposals mandatory. In the following paragraphs we will offer examples to illustrate the diversity of SEA actions of the "programmatic review and recommendation" type.

One western state adopted a "preferred way to go" with respect to the reading program, which was based primarily on the beliefs of the state Title I director informed by the Title I performance data. We were told, "We believe that a certified reading specialist is the best way to go in Title I reading programs in this state. If Title I LEAs do not have certified reading programs, then there must be a very close relationship between the Title I program and the regular certified reading program." The SEA representatives did not demand this program orientation, but they did "try to impress on LEAs that these are appropriate ways to go."

Another western state uses supplemental Title I data to make financial allocation decisions and mandate certain instructional arrangements. The require LEAs to submit expenditure data in addition to the standard TIERS performance information. This financial information is used in a number of ways. In addition to use of the detailed cost data as a guide to distribution of monies in the state, the cost data are reviewed for their instructional implications.

An eastern state takes a more indirect approach. They are concerned about the extent to which the LEAs "attend" to the recommendations they themselves have made as part of the evaluation report in prior years. (Such recommendations are part of an additional evaluation report required of Title I LEAs by this particular state.) To encourage LEAs to pay attention to their recommendations, the SEA requires them to indicate what they have done about these recommendations in an annual interim report. In addition, when monitors from the state visit each LEA they inquire whether the recommendations "have been attended to, not necessarily done, but attended to. Have the project director and other administrators given them reasonable consideration?"

This state also uses the additional evaluation data reported by each LEA to make informal judgments about effective instructional strategies, and they convey this information to districts. "In more successful Title I programs we find that these kind of things seem to be happening--these kinds of factors might reasonably be tied to successful programs for other students."

Other states have used the data from Title I evaluations as the basis for recommending model programs for LEAs to emulate. A southern state identified exemplary programs based on their achievement scores--programs that had an aggregate NCE gain of at least seven--and made a presentation describing each program at an annual education fair attended by representatives of every LEA. It was hoped that the LEAs would adopt similar programs themselves. Nothing was done to require LEAs to attend to these programs; however, there was every indication that they do take notice. In the past several years visitors to the fair have been asked to "identify the single display most educationally meaningful and most beneficial if adopted by other districts" and the exemplary Title I projects have been mentioned most prominently.

A midwestern state was about to begin a similar activity when we interviewed there. With TAC assistance they were "in the process of developing the conceptual and operational plans necessary to identify model Title I programs around the state and to share this information with all Title I projects."

A midwestern state "found that kids were performing at a higher rate in pull out programs than in regular classrooms" and ended up "advocating that change (in cases where racial integration was not a factor)."

An eastern SEA analyzed Title I needs assessment data across the state in an effort to determine the strengths and weaknesses of Title I programs statewide. In order to make a valid comparison the state "designated the same needs assessment procedure to be used by all LEAs with the state paying for the cost of testing on the CTBS for

every child at grades 3, 6, and 9." Based upon systematic analysis of these data they determined that there had been "considerable improvement in reading, but that math was in trouble." As a result they acted to "encourage and foster a change towards an increased amount of effort in mathematics." The data allowed them to be quite specific in their recommendations because they "could point out in evaluation reports what grade levels the most substantial gains were being made in existing mathematics programs." This information was communicated to all LEAs in reports issued by the SEA.

This lengthy listing of instances in which SEAs used Title I data for "program review and recommendation" illustrates the diversity of such activities. They were the most frequent examples of SEA evaluation use we found in our interviews.

Input into State Decision Making. There were also a number of examples of SEAs using the data from Title I evaluations as input into educational decisions being made at the state level. For example, one state used the annual Title I evaluation report "to justify the state funded compensatory program--the needs data were used to justify state resources by showing the numbers of unserved children and by demonstrating that the program can produce gains." In this way, the Title I report "presented knowledge of gains and operations of basic skills which provided information on how the state should set up its own programs for non-disadvantaged."

Another state also used the data from the Title I evaluation as a yardstick for measuring the success of its own funded programs. This midwestern state compared the LEA achievement scores in Title I with the achievement scores in the state funded compensatory program.

An eastern state called upon the expertise of Title I staff to provide input into various educational decisions at the state level. The Title I staff were often asked to meet with various task forces and present evaluation data. Title I staff were frequently able to use their evaluation data to make technical recommendations to other programs. One such activity related to "educational competencies and learner outcomes" and the data were used to assist in "designating specific grade levels for different competencies."

A midwestern SEA consulted with Title I evaluation staff during the process of developing evaluation systems for its special education and vocational education programs. Thus, the Title I evaluation was used to help the SEA develop other evaluation systems.

In summary, we found many examples of SEA use of Title I evaluation data. All states monitored LEA compliance with Title I regulations. In addition, almost all used TIERS data in conjunction with other supplemental Title I evaluation data to help recognize problems and/or identify exemplary programs. Sometimes the SEA role ended when the data were communicated to LEAs, but in the majority of cases SEAs did more. Most provided monitors who worked with LEAs on problem areas. Beyond this, more than half of the SEAs had used Title I evaluation in problem solutions of one kind of another. These actions involved deciding on technical assistance strategies, making program recommendations to LEAs and using data in state level educational planning and decision making.

The Influence of TACs on SEA Evaluation Use

As a part of the SEA interview, we sought to determine the ways that TACs had influenced evaluation use in SEAs. As we noted earlier, the limited research on TAC effectiveness suggests that they have been of most help to agencies when they provide four types of service: fostering understanding and acceptance of evaluation generally and TIERS specifically; improving clients' abilities in testing and reporting procedures and data interpretation; providing "how to use it" workshops; and helping with the development of other evaluation procedures to meet local needs. Our interviews with SEA personnel confirm these suggestions. SEAs enthusiastically cited examples of TAC assistance which cut across the four categories.

Several SEAs extolled the role played by TACs in shaping their attitudes toward evaluation. One southern SEA commented, "The TAC has been very helpful in getting us to think about evaluation use, and in increasing evaluation use." They felt that TACs had helped them and LEAs to understand the quality of TIERS evaluation data, which in turn made LEAs more willing to use it. As they said, "The LEAs have become more aware of evaluation and do a better job of evaluating and want to use it." In the same vein, an eastern SEA remarked, "The TAC has helped a whole lot in helping the LEAs to understand evaluation and its use in decisionmaking. There is more interest in providing and using information for program improvement--a big difference from 10 years ago."

SEAs gave high marks to TAC assistance in improving testing procedures and data interpretation. In one midwestern state, for example,

the TAC worked with the staff in providing inservice training on the meaning of evaluation data. The SEA commented, "The TAC assistance has been fantastic." In another instance, TACs in one southern state conducted workshops on the match between tests and curriculum. The SEA was impressed because, "As a result of the TAC workshops, some of the LEAs have changed their tests."

TAC sponsored workshops were also highly regarded by our SEA respondents. One SEA, for example, stated, "There has been a very heavy involvement of TACs in the conduct of workshops about TIERS models." A western SEA asserted, "The TAC Centers were a very important part of the inservice training provided to the staff of the state office as well as to the LEAs." In this case, the relationship begun in the workshops continued through the evaluation cycle with TAC personnel accompanying SEA evaluators on their annual site visits to LEAs.

Involvement in SEA site visits is an example of TACs helping establish additional evaluation procedures for specific local needs. A number of instances of this type of assistance were mentioned by SEA personnel. One midwestern SEA presents a typical set of comments. "They have done additional kinds of things with respect to interpretation and analysis of evaluation data. They have also implemented studies to spot check data submitted to the SEA by districts."

From our interviews with SEA personnel, it is clear that TACs play a part in SEA evaluation use. They were mentioned often by SEAs, and almost all of the comments were positive.

CHAPTER 4

TITLE I EVALUATION ACTIVITIES AND USE -- LEA LEVEL

Introduction

Our analysis of the utility of Title I evaluation at the LEA level is based on two major sources of information. The first, a review of previous research, was presented in Chapter 2. The second data base consisted of first-hand reports of Title I evaluation utilization at the LEA level derived from interviews with parents and staff in 12 LEAs. This information will be presented in the current chapter.

Both sources of information suggest similar conclusions about the utility of Title I evaluation, i.e. its potential for local use. The literature review indicated that Title I evaluation did possess local utility. While the reporting and compliance functions of Title I evaluation predominated in most earlier studies, there was also evidence that Title I evaluation data were useful at the local level. Moreover, the review indicated a trend toward increased local evaluation use in the past couple of years. This coincided with increased emphasis on local utilization of Title I evaluation on the part of the Title I Technical Assistance Centers.

Because of these recent shifts in emphasis, an accurate appraisal of the current utility of Title I evaluation required an assessment of the situation at the present time. Consequently, the bulk of our efforts were directed towards interviews and case studies in selected LEAs.

We found widespread evidence of local use of Title I evaluation in the 12 LEAs we examined. LEA personnel reported that they used Title I evaluation in a number of different ways: it helped them recognize situations that needed attention; it was a force for changing attitudes and opinions about educational issues; and it fed into decisions regarding organizational and instructional change.

As this brief capsule summary suggests, we did not uncover dramatic instances in which evaluation data acting as the major force in a significant program decision. Instead, we found a multiplicity of situations in which Title I evaluation played a part in small actions and in gradual, incremental program evolution on a larger scale. There was consistent evidence that Title I evaluation was used by the LEAs in our sample. This data, which will be elaborated in the present chapter, argued strongly for the local utility of Title I evaluation.

Our analysis of LEA evaluation utilization went further than merely compiling and categorizing instances of use (though this was important in its own right). We conducted in-depth case studies in four LEAs. These case studies (which are included as appendices to this report) provided additional insight into some of the factors that affected Title I evaluation use at the LEA level. In addition, they contained descriptions of some of the more impressive instances of evaluation use. These selected cases suggest the great potential that exists for utilizing Title I evaluation at the local level.

This chapter is organized in three sections. In the first we summarize the various instances of LEA evaluation utilization in tabular form, and present some descriptive statistics. Following that, we offer

capsule descriptions of LEA evaluation use drawn from the 12 districts we examined, and we discuss some of the underlying relationships that emerged from our investigation. In the final section we draw upon the qualitative data from the case studies to speculate on factors that affect Title I evaluation utilization at the LEA level.

Descriptive Summary

Based on recommendations from TACs and SEAs and an initial set of screening interviews, we identified 12 districts that were believed to be models of effective evaluation use. We solicited descriptions of instances of evaluation use in conversations with personnel from each of these LEAs. Table I presents the instances of evaluation use mentioned in the field interviews, organized by type of utilization and type of evaluation information. In brackets, we have recorded duplicated instances. Thus, 4(7) refers to four different events, some of which were mentioned more than once. Results discussed in this report are based on the single incident tabulations.

Evaluation Information Use: Information from the Title I Performance Report and from TIERS was the evaluative data type most frequently mentioned. Specifically:

> 43 percent of the instances of evaluation use described to us were impacted by information from TIERS (both the Title I Performance Report and the TIERS "Models" data).
>
> 19 percent used data derived from local process evaluations.

Type of Utilization	Type of Evaluation Information: Needs Assessment	TIERS	Local Evaluation	Student Process	Other Selection	Other LEA Testing	Total
Attitude Change	4(4)	18(15)	10(8)	1(1)	12(9)	4(4)	49(41)
Situation Recognition	2(2)	28(24)	7(5)	2(2)	4(3)	8(7)	51(43)
Administrative or Organizational Action	4(4)	20(18)	11(10)	8(6)	7(6)	4(4)	54(48)
Curriculum or Instructional Action	2(2)	29(20)	11(11)	3(3)	10(7)	5(3)	60(46)
Total	12(12)	95(77)	39(34)	14(12)	33(25)	21(18)	214(178)

TABLE I

Instances of Evaluation Use in Exemplary Districts

Number of Incidents (Unduplicated)

14 percent used data derived from other LEA testing.

07 percent used data derived from needs assessment, while the remaining 17 percent were divided evenly between data derived from testing for student selection and various other types of evaluation.

The category "Other" included any evaluation technique which could not be considered part of the other data sources. Most of the incidents in this category involved personal judgments made on the basis of observation. There were some novel examples, however, including the use of photographs in a preschool program and monthly teacher/tutor conferences on students.

Type of Use: We categorized the various uses described to us by respondents into four groupings: attitude change, situation recognition, administrative or organizational action, and curriculum or instructional action. In examining our data we found that while all four types of evaluation use were well represented, evaluation was used most often as part of an action to modify administrative or organizational patterns. Specifically:

27 percent of the incidents cited involved administrative or organizational changes.

26 percent of the cases involved curricular instructional actions.

24 percent involved recognition that a situation existed which required attention.

23 percent involved changes in attitudes.

Typically, the uses of Title I Performance Reports and TIERS for attitude change involved attempts to persuade others about the value of the Title I program. Examples of this phenomenon include:

> A Title I Coordinator in an eastern district who used TIERS pre-post achievement data to demonstrate the success of the Title I program to members of the school board. She hoped that increased local funding could be obtained to offset anticipated declines in federal support.
>
> A Title I Director in another eastern LEA who used summary TIERS statistics in his presentation to teachers to help them formulate more reasonable expectations for student performance.

In most instances the use of Title I Performance Reports and TIERS for situation recognition involved the discovery by the Title I Coordinator or Title I Director of previously unknown facts about the program. Examples of this situation include:

> A Title I Coordinator in a southern district who discovered a surprisingly high rate of absenteeism when he reviewed the Title I Performance Report. He communicated this information to the Title I Director, who had previously not recognized the severity of the problem.
>
> A Title I Coordinator in a midwestern LEA who formulated TIERS data for use by building principals so they could diagnose their own problems. As he described his attitude about problem recognition, "You can't have a knee-jerk reaction (to evaluation data).

You can't tell building principals to do this or that, but you can change what they think of this or that--you want reflective thinking."

Teachers and parents in a southern district who jointly reviewed students' performance scores to better diagnose problems so they might consider ways for parents to assist with their child's education.

The following examples are typical of the kinds of administrative and organizational changes that are informed by TIERS evaluation data.

A Title I Coordinator in a midwestern district who was prompted by TIERS reading scores to examine the operation of both the Title I reading program and the regular reading program by carefully observing classes. He then took administrative actions to better articulate the two programs.

An assistant superintendent in a western district who used the data from TIERS to make administrative decisions on program cuts. As he put it, "when you have an expanding program you can get by with limited evaluation data, but when you have to make cuts the data become very valuable."

16 percent of the districts received direct assistance in interpreting test scores.

School Board and Various External Agencies

We chose to include the School Board along with various agencies external to the school as a distinct category. In the way that TIERS and other evaluation information are presented to (and used by) them, Boards seemed to be much more akin to SEAs, joint dissemination review teams, and districtwide parent advisory committees.

In a significant number of instances, school boards and other agencies actively sought out information for decisionmaking. A board member in one southern LEA told us that Title I test results were very important in the board's deliberations about district staffing for the following year. "The board authorizes every position annually," he explained, "and the Title I results were considered along with the Title I Director's recommendations in determining how many positions would be funded." In a small western LEA the School Advisory Council in the course of approving Title I funding allocations each year, sought out and used annual TIERS test data and Title I needs assessment data. As the Title I evaluator told us, "The reason we (the evaluation department) have been around so long (nine years) is because we give the local School Advisory Committee something they can use."

The annual Title I performance data collected as a part of TIERS was the information that seemed most relevant to the school board and to other external agencies. This seems sensible. Agencies that have broad advisory capacity are more likely to be interested in summative data that capture the status of a program over a long period of time. The TIERS reporting system is quite appropriate to this need, and it was not surprising that most of the evaluation that was used by these

agencies came from TIERS. Sometimes, however, the TIERS data were not so much used by Boards as by administrators to influence Boards.

In many of the discussions involving use by Boards of Education or other external agencies, the district administration presented evaluation data to persuade or convince the external agency of a particular position. Our respondents used words like "persuade", "justify", and "convince" when when they discussed these instances of use. In one eastern LEA the Title I Coordinator indicated that TIERS data were used to help persuade school board members that the Title I program was worthy of increased local support. The coordinator anticipated a decline in federal funding and hoped to use the positive test data to generate additional funds from local sources. Many of our respondents mentioned similar activities that were designed to generate "local support" from external groups, though not all of them had such strong overtones of coercion.

In many cases "generating local support" meant keeping parents and community members informed of the successes of the program. In one large eastern district the evaluator described how longitudinal TIERS data had been presented to the Parent Advisory Committee to indicate the success of Title I. He was pleased to be able to say, "In 1972 we were 12 percentage points behind the average. By 1981 we had cut that gap down." Such positive reports keep the morale of the parents and the local community high; this is important in improving staff morale and keeping individual students motivated.

The use of TIERS data to influence Boards or other external agencies should not be viewed as negative in tone. If information is posi-

tive, administrators rightly should reap the rewards. Thus, for example, a number of instances were reported in which Joint Dissemination Review Panels were provided with TIERS data as part of a presentation about the success of the program. Similarly, in one eastern school district, TIERS data was presented to the state Human Relations Commission to persuade them to relax the rigidity of certain state requirements and provide a greater amount of time to implement a certain plan. The superintendent who spoke with us felt that Title I success, as evidenced by test performance, was an important factor in winning the Commission's approval of a relaxed time schedule.

This is not to suggest that TIERS data were the only type of evaluation that was brought to such agencies. In a western state the Title I Director reported that he used both TIERS data and the Title I enrollment data in presentations before the non-public school advisory councils and non-public school administrators. "We have a fairly large non-public system and the data helped me justify to a large number of people that we are meeting legal mandates. It is essentially a public relations service, but we also determine personnel allocations on the basis of these data."

To summarize, TIERS and other evaluation data were reviewed frequently by school boards and other external agencies. They were used to make summative judgments about the success of the program which then formed the basis for various types of action. In most cases the contact was initiated by a district administrator who was seeking some particular response from the board or the advisory committee. In a few cases, however, the board reviewed the data at their own initia-

tive as part of their deliberations surrounding an administrative action. Annual achievement test data such as that reported in TIERS were most useful in these situations.

District Administrative Uses

We categorized district administrative uses as pertaining to broad policy decisions that typically affected more than one individual school site. Individuals involved in such uses frequently included the school superintendent and deputies, curriculum specialists, the districtwide Title I Coordinator or federal projects coordinator, and the district evaluation staff or Title I evaluator.

A general perception is perhaps first in order. At the level of the district administration we observed somewhat fewer references to TIERS data per se and more references to locally-developed Title I evaluation. In addition, most of the uses at this level were direct--related to decisions made by the individual rather than to influencing or persuading others' attitudes or actions. Typically, the superintendent, the Title I Coordinator, or the administrative cabinet supplemented their own impressions with TIERS data and other local evaluation information to make districtwide program decisions.

In a large eastern district the evaluator conducted an extensive examination of prior years' TIERS data to determine the relationship between achievement and the frequency with which students were served in the Title I program. The results of this evaluation were used by the Director of Specially Funded Programs and the Superintendent to establish program guidelines for frequency of Title I contact at various grade levels.

Another noteworthy examples comes from a smaller western district. The Title I evaluator told us how the information from norm-referenced testing (TIERS) and district criterion-referenced testing (CRTs) was useful at different levels in the administrative heirarchy. In particular the TIERS data, administered from year to year, helps the LEA see districtwide patterns among the Title I schools and has impact on program decisions. The coordinator asserted also that many district decisions on the activities conducted in the Title I program and the program emphasis were based in part on these evaluation data.

In a small southern district a principal described how the Title I Coordinator used TIERS results longitudinally to examine the Title I delivery system. The coordinator concluded that the program was more effective with small numbers of students and at younger grades. As a consequence, the district initiated a preschool summer program for small groups of students who would be entering kindergarten. Another evaluation conducted in this same southern district compared three alternative instructional programs including a computer assisted instructional option. The evaluation used TIERS along with other evaluation data collected under the auspices of the Title I project. The results of the evaluation showed that the computer based program was as effective as the other two programs and had additional advantages as well. Consequently, the district approved the purchase of additional computer equipment.

There were many instances in which the TIERS data considered at the district level were used not so much for specific program decisions as for program recognition. For example, the assistant superintendent

in a western district indicated to us that she used the Title I evaluation reports to identify 10 to 12 schools per year that were to be targeted for special assistance. In this way she was able to focus limited funds and resources on the schools that had the greatest need. Her judgment was based on the evaluation reports that she received from each school. In fact, she was so pleased with the usefulness of the evaluation that she told us, "If the external funding was stopped, I'm convinced we would continue our evaluation anyway, it has been that helpful."

In one western district the Title I Coordinator explained how he used the evaluation data to prepare consultants prior to visiting the schools. This made their monitoring efforts more efficient since they had some familiarity with the school's program in advance.

Use of supplemental evaluation information was more prevalent at the district administrative level than at the board level previously discussed. For example, in one small eastern LEA, information from a special "assessment of distribution of Title I services" (which was an implementation/process evaluation undertaken as part of the coordinator's participation in a TAC workshop) provided administrators with a better understanding of the types of services that were being offered to Title I students. The data confirmed the administrator's perceptions about the program and contributed to the decision to continue in the direction they had been going. "The special evaluation was designed to answer the question, 'what are we doing?', it gave us better information about what the tutors were doing."

The Title I evaluator in another eastern district described an instance in which the LEA hired an outside consultant to visit classes

and prepare an evaluation report on the way the Title I program was organized at a particular school. The report was highly critical and the district sought assistance from a second expert. When the second report confirmed deficiencies in the process, the school made changes in the program.

Frequently, it was not a specific evaluation report or test score that had impact but the process of interacting with the Title I funded evaluator. One of the best examples of this came from a western district in which the evaluation branch and the subject matter consultants worked together extremely well. The Title I Coordinator told us, "The evaluator and I wear each other's socks, so closely intertwined has become evaluation and instructional inquiry." They reported that the availability of good Title I evaluation had caused them to develop a more questioning attitude and had provided a means for exploring a number of educational questions. One example in this district was the initiation of a specific evaluation study (at the request of the Title I Coordinator) to provide answers to questions about the relationship between pupil maturity and teaching methods.

The assistant superintendent in another western district explained that parent surveys were particularly useful in helping the district to identify problems. This data source was part of the Title I evaluation effort.

The impact of Title I evaluation on improved instruction and operation in various non-Title I aspects of districts was noted on numerous occasions. The assistant superintendent in a southern district told us how he was able to make use of Title I evaluation requirements to

accomplish things he could not otherwise have done. The demographic data collected as part of Title I evaluation proved to be quite useful when he had to "set a profile for the district" for other grants that he was writing. As he told us, "The Title I mandate makes it easier to get these surveys done."

Our intensive interviews in local districts revealed a sizeable number of unanticipated consequences of evaluation that our respondents found to be particularly "useful." The Title I evaluator in a western LEA explained that the district evaluation office (which is separate from the Title I evaluation office) was becoming more sophisticated as a consequence of the Title I evaluations and the TIERS system. "We (Titie I) have had to be more sophisticated because we have had to please the federal demands. This has, in turn, upgraded the quality of the district evaluation staff," the evaluator explained. Though this is not an intended use of Title I evaluation, the beneficial effects of upgrading school staff capacity is, nonetheless, a pleasant side effect.

In an eastern district the director of funded programs decribed how group interviews among Title I teachers had been held at the district headquarters as part of a local Title I process evaluation. The purpose of the interviews had been to provide data on teachers' perceptions of the success of the program. More importantly, the process brought teachers together in an atmosphere where they could talk about the Title I program implementation and share information. This proved to be particularly useful to teachers in the TESL Program (Teaching English as a Second Language). A positive side effect of the Title I process evaluation was increased communication among TESL and Title I

teachers. From this communication the TESL program developed a more structured curriculum and a better sense of program direction.

Some evaluation uses occur at the building level and decisions made, attitude changes, etc. affect either all of the classrooms in the school or some significant subset. We have made the distinction between uses at the building level which affect the whole school, and uses at the classroom level which affect a single group of students and a single teacher. Building-level uses can involve the principal, other administrative staff persons, resource teachers, the rest of the instructional staff, and parent advisory committees. Some of the actions described to us were taken by the administrative staff alone (many involved only the school principal). Yet, there were other cases in which the entire school staff was involved.

In many LEAs Title I data were directed back to the individual schools with a clear purpose in mind. The Title I Director in a western LEA explained to us that he believed very strongly in "data-based decisionmaking." There had been a lot of effort in this district to have the individual school principals focus more of their attention on the Title I evaluations when they developed their program. The district evaluator felt that principals had become more interested in data and found it more useful in decisionmaking as the emphasis on data-based decisionmaking had grown over the last few years. We suspect that many principals had been "encouraged" to make use of the Title I data in program decisionmaking. Nonetheless, our interview with a principal in that district provided evidence of school level use of the TIERS and other evaluation data.

Another example of direct district pressure on a school principal came from the Title I evaluator in a midwestern school district. He confronted the principal in one school with a comparison between that school's pre-post TIERS scores and the district averages. The school's profile was below the overall district profile in the district. The principal acted quite defensively; he was particularly defensive about a learning machine that was being used as part of the school's program. However, the weight of the data finally convinced the principal to look at other alternatives, and a new learning machine was selected. As the Title I district evaluator told us, "The impetus was the evaluation results."

Building-Level Uses

The specific instances of Title I evaluation use at the building level described to us focused on the standardized test scores reported annually as part of TIERS. A southern principal described to us how a committee of parents, teachers, and administrators planned the school's program each spring and used the previous year's standardized test scores to select programs.

Several additional examples may prove elucidative. In a southern school district, a principal explained how he personally made a decision to switch reading programs when the standardized reading test scores indicated that a change was appropriate. A principal in a western LEA provided a similar description of program planning decisions made jointly among teachers and administrators based on the pre-post test scores that were part of TIERS. Finally, a Title I Director in a

western school district explained how she met jointly with teachers and the school principal to review the TIERS data and jointly decide on ways in which the program might be changed in the coming year. "The reports form the basis of our discussions--and their decisions," she told us.

There were a number of instances in which the LEA broke TIERS data down into their own format before it was reported to the schools. Such efforts were well received because the information was more readily useful. At least that is the way the principal in a western school district described it. "I can't tell you how important those data have been to us. A key to this has been the willingness of the evaluation people to answer dumb questions and not to make us dumb in asking them." The district evaluator reformats the data before presenting it to the schools. A school informant indicated that the data were used to identify problem areas, select materials, and develop teaching strategies.

A principal in a small southern LEA told us that she felt the CTBS scores were the most important piece of data she had for decision-making. She explained that Title I selection was based on these scores. "CTBS scores are the most important piece of information and in a sense measure the quality kids in Title I." A principal in a small eastern school district also reminded us that one of the important uses of Title I evaluation at the school was for selection of students, but in his case the staff used a wide variety of instruments.

Data other than from tests were viewed by participants as useful. A teacher in an eastern district told us that she used information about

the recidivism rate drawn from a Title I evaluation report to persuade parents to enroll their child in the Title I program. She reminded us that "you really have to sell some parents", and "a statistic like that" is useful in convincing them of the benefits of the program.

Just as the evaluation process had occasionally had impact at the district level sometimes such impact was felt at the school level. An assistant superintendent described the situation. It was not the evaluation results that had an impact, instead the administrator was pleased because the evaluation process had an impact. As he described it, the evaluation helped the entire Title I project at certain schools become more "targeted." According to this district administrator, the teachers and the staff developed better definition for their project, developed measurable objectives, and were conducting the instructional program more effectively because they had to develop an evaluation strategy. "Evaluation helps make this process more definite, makes us more sensitive to specifics, to management..."

Classroom-Level Uses

A wide variety of tests were useful at the classroom level. Not only TIERS data but data from many supplemental Title I evaluation and testing efforts were mentioned when respondents described these classroom-level uses. In most of these cases, teachers took the lead in incorporating available information into their instructional program. However, in some instances, classroom uses were described in which someone else was using evaluation to convince a classroom teacher to behave in a particular manner. We will consider a few of these types in turn.

In a large midwestern district the Title I Evaluator explained that the district recalculated TIERS test scores into six clusters of items that have curricular relevance for the district Title I program. This information was provided to individual teachers whose response was very positive. This kind of use was fairly typical of a number of instances.

A teacher in a western LEA told us that Title I test data were very useful for placement, but that they were used in conjunction with her own knowledge of the students. "It (the data) sometimes makes me take a second look at some of my instructional decisions. But sometimes the data (tests) are simply not an accurate picture because of good guessing or bad test taking." A teacher in a small eastern school district told us that test data of various kinds were useful to assess the child's performance data and recognize strengths and weaknesses. Summer school test data from the Title I program is extremely useful to regular school teachers in one southern LEA. This information helps the regular school teachers know children's standing at the beginning of the year. "From the test results from summer school I know what areas to start with. I feel that I am more efficient. It is a big help and advantage."

Teachers do not always rely on the TIERS standardized test data, but frequently supplement it by other more focused tests (criterion, references, skill charts, diagnostic, etc.). This "supplemental" testing was part of the Title I evaluation. One southern school district had developed a very elaborate reading skills chart in which partial mastery or non-mastery was indicated for each child based on the previous

year's test results. Teachers could structure their lesson to address the specific areas of the child's weakness. The junior high counselor labeled this "a testing program that plugs into the school program" but a teacher added, somewhat less enthusiastically, "the skills chart certainly makes teaching, I wouldn't say simpler, but at least I have a goal." A teacher in a large eastern district indicated how multiple diagnostic tests were used as part of Title I to determine student needs and to direct instruction. In this district the teachers used the test results to decide how to group the children and develop remedial strategies. The external evaluator of an eastern school district that used diagnostic testing extensively said it was useful in individual classrooms. "If we couldn't get the teachers to see some value in the testing process, then there really was no point to testing. We began to use tests then as tools for diagnosis. We were not interested in showing change as much as in using the tests to diagnose student needs." He felt that this diagnostic perspective had been extremely effective.

A secondary level Title I teacher in a western school district praised the use of multiple indicators. "People tend to blindly believe any test results. These multiple test data help us to gain a better picture of what we're doing and avoid simplistic mistakes."

Finally, we should note that evaluation, particularly Title I evaluation, was useful to inform parents of the progress of their children. A parent in a southern school district explained that test scores were made available each spring so parents could see "how much their child has learned year by year." In fact, one parent we interviewed indi-

cated that conferences with the teacher that focused on test scores helped her assist in her child's reading instruction at home. "I'm glad they tell me he's falling behind. Title I will help him and I will help him."

Evaluation data was also used by administrators and evaluators to try to convince teachers to conduct themselves in a particular manner. The Title I Director in an eastern LEA described how he reviewed test data annually and reported back to teachers in an attempt to help them set reasonable expectations for student growth. More directive was the Title I evaluator in another eastern LEA who felt that the imposition of five-times-a-year criterion-referenced testing "forced teachers to pay attention to student progress."

In a large eastern LEA, consultants were used to monitor classroom behavior. The consultants visited classrooms and made judgments about the success of certain instructional strategies. As a result of one consultant's comments, workshops were set up in which teachers were encouraged to circulate throughout the classroom interacting with students rather than managing instruction from their desks.

We want to relate the remarks of a Title I school coordinator in a southern school district who felt that the process of taking attendance at parent meetings had been beneficial because it indirectly "gives teachers an impetus to get parents involved." She felt the teachers were more likely to work for parental participation when they knew that records are being kept. The hidden implication, of course, was that these records were going to be used and teachers would be judged on the number of parents who showed up from a particular classroom.

Summary

In this section we reported a number of specific instances of LEA evaluation use drawn from over 100 instances that were described to us. These examples illustrated the fact that use can occur at many different levels in an LEA, and moreover, we demonstrated that different patterns of use predominate at different organizational levels in the school district.

To some extent there were also differences from level to level in the balance between the use of TIERS data and the use of other Title I evaluation data. Summative data, such as TIERS, dominated large policy decisions made by school boards, district advisory committees, and other external agencies. In many of these cases the administration used data to persuade the advisory groups to behave in a certain manner.

At the district level, on the other hand, TIERS data was mixed almost equally with other Title I evaluation developed within the district. At this level, most of the data was directly useful to a decision which the administrators were going to make themselves. At the building level the mixture of data types was similar though there was somewhat more reference to TIERS.

Finally, at the classroom level a wide variety of tests and evaluation information was useful. The more closely geared the information was to the instructional program the more likely the classroom teachers were to make use of it. There were a larger number of instances in which someone was using evaluation to influence one party (classroom teachers) to act in a particular manner than we observed at the building or district level.

CHAPTER 5

FACTORS AFFECTING UTILIZATION IN THE CASE STUDIES

Earlier chapters presented the broad conclusion that widespread but subtle use of evaluation results occurs and instances of specific and sometimes elaborate utilization were presented. At this point we will focus on a small portion of the data set--the LEA case studies (Appendices B-E) to draw conclusions about factors leading to utilization.

An examination of the qualitative data presented in the case studies leads to six major clusters of factors. These have been identified in other literature on evaluation use as well as in the more focused Title I related studies presented in Chapter 2. The first, the personal reputation and credibility of the evaluator, leads logically to the second influence --the commitment of the evaluator to utilization. The third factor cluster was the interest of decisionmakers and community in evaluation use. Local focus of evaluation was a strong fourth influence, while the way evaluation results are presented formed another powerful enhancement to utilization. Finally, we concluded that evaluation utilization was increased when decisionmakers got help in developing procedures to use evaluation data. The rest of this chapter looks at the case study examples which led to these conclusions.

Evaluator Reputation and Credibility

It should come as no surprise that the way the evaluator is perceived played a leading role in affecting evaluation utilization. Alkin et al. (1979) identified credibility as a relevant factor in evaluation use. Evidence from all of the case studies reiterates the importance of this influence.

The reputation and personality of the evaluator was consistently cited in the studies. King noted that the effective use of Title I evaluation information in Small Town was "a tribute to the skill and hard work of two women (the evaluators)", one described, "as everyone around her is quick to point out, a positive force in the local school community." About the other woman, King noted, "One teacher labeled her a 'unifying force' and the 'only source of information for our Title I Program.' The parents and principals also shared this feeling, suggesting the potent effect of a single person's commitment to the use of information in a school district." She went on the assert, "Her sense of purpose and positive attitude were the critical attributes she brought to the job."

Williams, talking about Suburb West's evaluation usage, described "the evaluation specialist everyone credits with developing Suburb West's evaluation system." Daillak found that in Big City, "Mr. Green undoubtedly helped persuade the Region's administrators of the value of evaluation," and continued, "Mr. Green's university training prepared him for the technical demands of his job, but his personal characteristics truly suited him to it." More than that, Daillak observed, "Mr. Green created an evaluation role in Region D that fit his own personal style."

According to Braskamp, the evaluation director in Middle City was respected in the district even though his information sometimes caused resentment. The assistant superintendent told Braskamp that "The evaluator is brutally honest. He wants the truth to be known."

King credited the facilitating role of Small Town's evaluator, Mrs. Foley, to the way she was involved with all levels of the school district. According to King, Mrs. Foley dealt with the counselors and principals at each school. In addition she worked with "the teachers, especially at the junior high school, helping them interpret test results and target their instruction accordingly." This was not the end of Mrs. Foley's involvement. "(Her) role generates the use of test results not just by teachers or even parents, but to the students themselves...her personal contact with the children who take the tests encourages them to pay attention to the results."

The importance of evaluator respect and credibility in securing evaluation utilization is perhaps best summed up in the Big City case study. Daillak observed of the evaluator, "Harold Green appeared to have carved a niche for evaluation in Region O."

Commitment of the Evaluator To Evaluation Use

The attitudes of the evaluator toward evaluation use have direct bearing on utilization. The evaluators in our case studies had strong beliefs in the importance of evaluation utilization. In Suburb West, for example, Williams found, "Mr. Leonard has a fundamental belief that evaluations are a waste of time unless they are used by someone." Small Town's evaluators began with a slightly different slant, "but

(their) attitude was a positive one, i.e. given that we have to do these things, how can we make them work for us?" As a result of this positive attitude, the evaluators "sought to make the experience worthwhile and valuable." King asserted that the result of this early positive attitude was that, "Both Mrs. Lopenia and Mrs. Foley want people to use the information they gather; the test scores are 'data worthy of being used and shared'".

The evaluators in the case studies were aware of the need to shape the attitudes of the potential users of their evaluations. The approach of Mr. Leonard in Suburb West is a case in point. Williams identified one of the basic premises guiding Leonard's evaluation program. "Negative data should not be feared. One learns from failures as well as successes. An atmosphere must be created so that advisory committees will not feel threatened by negative results but will, instead, view negative evaluations as a challenge to correct and improve current practice." Leonard worked to create positive attitudes toward evaluation and thereby increase its utilization. The effectiveness of his approach is illustrated by one of the changes in Suburb West which Williams describes. "In the past, school personnel feared data, believing that somehow it would hurt the school or be used against it. Slowly the schools have learned that data can be used to identify and solve problems." Similarly, Daillak noted that in Big City, Mr. Green met with union stewards for all of the schools. "At that meeting, he explained his 'ground rules' for evaluations--anonymity for individual teachers and schools, and independence for building principals." By taking this approach, Green was able to "secure the cooperation of the

Union." Teacher attitude toward evaluation became more positive and less fearful.

One evaluator in Middle City explained a recent shift in emphasis toward program improvement that has focused evaluators' attention on utilization. "In the last three to four years, the emphasis has changed to finding out more about what is going on in the programs. We have recently done less monitoring of the process and more on performance. To the teachers, we are viewed as a source of information."

The case studies demonstrated the importance of evaluators' commitment to utilization. Evaluators may have great personal influence within their districts, but if they are not concerned with use--only data collection --utilization is unlikely.

Interest in Evaluation by Decisionmakers and the Community

The attitudes of decisionmakers and the local community toward evaluation was a factor strongly related to the level of evaluation use. The importance of administrators' attitudes toward evaluation has been recognized. One of the Western LEAs commented on evaluation utilization in their district by saying, "Some schools use the data more than others. It seems to depend on the openness of the principals to data." Williams' analysis of Suburb West supports this contention. He noted that "All acknowledged that the principal plays a key role" in determining utilization and that "a critical factor is the principal's openness and interest in the program." He interviewed two Suburb West principals and found them, "using the evaluations and very enthusiastic about them."

The importance of community expectations and attitudes toward evaluation is a recurring theme in these case studies. Small Town's experience provides an example of the importance of shaping community attitudes toward evaluation. King credits Lopenia and Foley with creating the positive attitude of the Small Town community toward evaluation. "What began as a Title I requirement is now an accepted piece of district assessment...testing and the use of results are standard and valuable procedures."

Suburb West provides another strong example. As Williams avered, "Suburb West Unified School District is a district with a citizenry that expects to be involved in decision making." He illustrates the impact of community attitude with a description of one incident. At a particular school there was a lack of parent input to an evaluation survey. "The parent committee members felt the low return rate was because parents didn't really believe that anyone would give their opinions on attention. Once the parent committee member reported that the data were indeed being used, the percentage of survey returns increased."

Local Focus of Evaluation

The three factors discussed so far might be considered interpersonal. The fourth factor to emerge from the case studies shifts the emphasis to the mechanics of evaluations. The case studies strongly suggest utilization increases when evaluations are specifically designed to meet local needs. Williams states this clearly. "Title I evaluation utilization can be greatly increased if a process that specifically addresses the decision needs of local school site advisory committees is

developed and implemented." He points out that, "SWUSD has had this need in mind as it has designed, developed, and implemented an evaluation process almost entirely focused at the local school Title I advisory committee."

Suburb West worked in another way to make their evaluation plan meet local needs. They got permission to change their testing schedule so that "it fits into the planning cycle of school Title I advisory councils." This change allowed Suburb West to collect and analyze evaluation data and report results back to the advisory councils in time for their April-May decisionmaking sessions. In addition, data are collected at different times in different schools. "As much as possible each school is allowed to select the most convenient time." Not only that, but SWUSD makes an effort to insure the results are delivered when needed. "The schools indicate the dates they want the data back and the Research and Evaluation unit guarantees their delivery one week prior to that date." In this way, "By April each school advisory council has all the data needed to begin deliberating its next year's Title I budget."

Big City provides another instance of locally focused evaluation activities. Daillak notes that, "Over the years, Mr. Green has provided special services, on an as needed or as requested basis." He has provided analyses of issues raised by administrators and prepared evaluation plans for program proposals, for example. These special evaluation services have resulted in a great deal of utilization. Daillak cites an analysis of test score declines conducted by Green, at the request of district personnel, which resulted in grade level reorganiza-

tion. In another instance, Green's evaluation of frequency of Title I service resulted in a proposed change in policy. Daillak attributes the success of utilization in Big City in part to Green's timely response and sensitivity to local concerns. As he says, "Mr. Green did his best to respond promptly and usefully to administrative requests for information," and, "He was responsive to teacher and school sensitivities."

Effective Presentation of Results

The fifth factor common to these case studies also emphasized the mechanics of evaluations. In all of the studies, the method used to present evaluation results, or the format in which they were presented, played an important role in gaining utilization.

Williams describes the ways in which Suburb West's evaluation data were reformatted to increase their usefulness to decisionmakers. The data were both simplified and presented graphically. "The data, which represent approximately 50 pages of computer printout, are reduced to a mere three pages and are presented in three ways: 1. The information is presented graphically in a manner which 'tracks' particular groups of students as the progress through the grades; 2. The information is presented graphically in a manner which shows the performance at each grade level for the various years involved; and, 3. The profile, in a unique fashion, presents the gains realized from one year to the next for the various grade levels." As Williams concludes, in essence the district "has developed a reporting format which allows Title I Advisory Committee members to assess quickly the present status of Title I pupils in their schools on ITBS subscales." The use of a

"highly visual, nontechnical format" enables rapid interpretation and encourages comparisons. According to Williams, "information is directly usable in the Title I and state program planning processes."

Small Town's evaluation team also converts its evaluation data into graphic formats for presentation. King said that Mrs. Foley "draws charts and graphs to visually show the results and to 'get information in usable form'". According to King, Mrs. Foley finds these graphic formats "of value in making presentations."

Big City believes in simplification of evaluation results, too. Daillak said Mr. Green, "Describes his job as that of transforming mountains of data into useful information and he valued concise interpretations and reports." He testifies to the effectiveness of this approach in Big City. "It seems likely that this style won listeners among the personnel at the Region office and contributed to evaluation information use."

Besides turning evaluation data into graphic, simplified formats, several of the districts in the case studies increased utilization by putting tne results into forms designed especially for the needs of their constituencies. For example, in Suburb West, Williams reports that Mr. Leonard developed two unique evaluation reports. "The APP and the EPA form the core of the district's Title I evaluation program. They have been carefully developed so that they collect relevant achievement and attitudinal data and analyze and present the data in a highly visual and easily understood format." Middle City has also developed a way of presenting evaluation results to its clientele. Braskamp notes that, "One of the first things that showed teachers the value of the evaluation department was its help in organizing reading test questions and

answers so they were of use in planning instruction." The evaluation department rearranges test score data and reports them to teachers in clusters of reading skills. The new presentation mode has been well received by teachers, and has had a fortuitous spin off. Braskamp quotes an elementary reading coordinator who said, "The cluster scores have helped the teachers' attitudes toward evaluation as well."

In Small Town, the evaluation results are used to create the Small Town Skills Charts. The Skills Charts present the results for each teacher's students using an objectives management system. King says of the charts, "The test results have been transformed into a highly explicit diagnostic tool for the Title I classroom, and the teachers are able to devise what amounts to an individual educational plan for each of their students."

Big City, too, has developed a unique format to present its evaluation results. Daillak describes the format by saying, "Region D's evaluation reports incorporated both test score results--the standard evaluation data required by funders--and narrative evaluation critiques." The narratives were prepared by subject-matter consultants and edited by the evaluator for inclusion in the final evaluation report. These critiques described program strengths and weaknesses and gave recommendations for program change. Daillak's evaluation of the narrative critiques indicates that they "appear to have gained a sympathetic hearing." In fact, the narrative critiques "appeared to be more influential than the test score summaries in the evaluation reports. The narratives made sense to administrators and teachers, and seemed to carry weight."

King, perhaps, sums up the effect of this factor on evaluation utilization. She said, "Written reports, memos, slide presentations, conversations, workshops--all are used on occasion to communicate results. Users are, in a sense, targeted, then presented information in a form appropriate to them personally." The use of effective presentation, she concludes, is one of the factors contributing to the use of Title I information in Small Town. The same appears to be true in Big City, Suburb West, and Middle City.

Assistance in Developing Use Procedures

The final factor to be discussed is an interesting example of a contributing factor. Suburb West, Middle City, and Small Town had evaluators who developed procedures to help their decisionmakers learn how to use evaluation data.

Mr. Leonard, Suburb West's evaluator, believed that "school administrators, teachers, and parents who serve on Title I advisory committees often do not have group process skills and decisionmaking skills. They must be given assistance in how to read, analyze, and make decisions upon evaluation data." To this end, Leonard has outlined a planning and decisionmaking process which lists the specific steps in sequence, the data that are available to inform the decisions, and the specific decisions that must be made. The special forms which have been developed by the district for reporting decisions have become increasingly familiar and useful to school advisory committees. Williams reports that Leonard and his staff spent a lot of time in the schools working to put this process in place, "Because the data-informed method of decision-

making was new to both school staffs and parents." Over the course of time, however, this process has become well established and "now that most school advisory committees are familiar with the techniques, the Research and Evaluation staff finds this consulting activity less time-consuming." Braskamp noted the effort Middle City evaluation department put into developing procedures to increase utilization. Middle City's evaluation dissemination system, "has produced a communication network that has involved all staff from Title I teachers to top administrators in the evaluation process." The dissemination procedure follows definite steps which ensure that evaluation results are passed to all decision levels, reacted to, and fed back through the system.

Small Town's evaluators have a slightly different approach. According to King, Mrs. Lopenia regularly encourages others to use evaluation data by giving them information and a related task or problem such as "suggesting changes in the program, planning next year's workshop, developing a dissemination plan, and creating a meaningful attendance policy." She is in essence training Small Town's personnel to use data in making school decisions by giving them practice with the process.

Summary

The case studies represent wide variations in geography, size, and organizational structure. Nevertheless, all four present clear examples of evaluation utilization. We hoped to discover common patterns among the variations which could lead to generalizable guidelines for increasing utilization.

We found six factors which contributed to evaluation use in all of the case studies. The first factor was the reputation and credibility of the evaluator. In every case, the evaluators had personal qualities which made them respected and well accepted by the decisionmakers. Their styles differed, but were appropriate to the milieu in which they functioned.

The second common factor was the attitude of the evaluators toward evaluation usage. All of the evaluators felt a commitment to seeing evaluation data used for local decisionmaking and took steps to ensure it happened.

The interest of school decisionmakers and the community played an important part in promoting utilization. The evaluators realized this, and took steps to mold positive attitudes. Their efforts paid off, and in each situation, decisionmakers expressed support for evaluation.

Evaluation data were used when they focused on issues important to decisionmakers--when they were collected to answer specific questions and meet local needs. Related to the evaluation focus was the fifth factor - the way in which evaluation results were presented. Each evaluator had devised presentation strategies which made the evaluation results clear, understandable, and useful to their audiences.

Evaluators who provided decisionmakers with step-by-step procedures increased utilization. The procedures varied from site to site, but they all gave guidance on how evaluation data could be used to make informed decisions.

The six factors we found in the case studies indicate specific steps which can be taken to increase evaluation use. They are factors which

are effective in urban, rural, and suburban school districts. We encourage people concerned with evaluation use to consider these factors carefully and to develop strategies which will ensure their presence.

CHAPTER 6

CONCLUSIONS AND SUGGESTIONS FOR FUTURE ACTION

Our review of the literature on Title I evaluation indicated a steadily growing awareness that evaluation utilization occurs. Moreover, the recent literature has begun to highlight the gradual incremental ways in which this utilization takes place. The SEA and LEA data we gathered in this study confirm and further amplify these prior findings. Both TIERS data and other types of Title I evaluation data are used at all decision levels. The Title I evaluation system does, indeed, have utility.

While we found relatively few examples of immediate, dramatic impact on decisionmakers, we uncovered strong evidence that evaluation data act as a contributing information source in the daily life cycle of projects and in incremental ways to more major program decisions. At the SEA level, evaluation data are used to monitor LEA compliance, to recognize both problem areas and exemplary programs, and to influence administrative and curricular actions. LEAs typically use Title I evaluation data to change attitudes and opinions toward Title I projects, to recognize situations requiring attention, and to contribute to decisions on administrative and curricular actions.

We also found that the different kinds of evaluation data have relative utility at the various organizational levels. School boards, district advisory committees, and external agencies rely on TIERS data more extensively than other evaluation data. At the district administrative

level, TIERS data are mixed equally with other Title I evaluation data developed by the district. At the building level, principals, coordinators, and the like rely slightly more on TIERS than on other data. Finally, at the classroom level TIERS data are rarely used; instead, data more closely related to the instructional programs are preferred.

Since we had selected SEAs and LEAs thought to have high levels of evaluation use, the identification of factors associated with that use constituted a primary focus of the study. Analysis of the case studies showed that six contextual variables affect evaluation's use. These factors are described below.

Evaluator credibility. The reputation and credibility of the evaluators in each of the case studies was an important determinant of utility. While each of the evaluators achieved credibility in differing ways, they all were perceived as competent and trustworthy.

Evaluator commitment to utilization. Credibility, while important, is not enough to insure evaluation utilization. The evaluator must also have a commitment to seeing that evaluation results are used by decisionmakers.

Interest in evaluation by decisionmakers and the community. Evaluation data are used when they provide information tailored to the needs and interests of the local school community. Utilization occurs when evaluators draw relevant information from TIERS data and when they conduct special evaluations to meet local requests.

Effective presentation of results. Graphic, narrative, and nontechnical modes of presentation increase the utility of evaluation data to local decisionmakers.

Assistance in developing procedures for the use of evaluation data. Evaluation use increases when decisionmakers are assisted in understanding how they might use the evaluation data. Successful evaluators in our study typically provided detailed, step-by-step procedures to potential users.

The results of this study lead us to suggest several recommendations for improving evaluation utility as Title I becomes replaced by Chapter

I in the new law. Chapter I evaluation utility may be enhanced by continued technical assistance in reformatting TIERS-like results to meet LEA information needs. Moreover, SEA and LEA evaluation units should be encouraged to design a variety of local decision focused evaluation strategies. In particular, locally designed evaluation procedures might provide information on the impact and costs of various materials and processes within projects.

Many local and state agency personnel require guidance in developing procedures to follow when making decisions. It is not so much that administrators do not want to use relevant information as they typically do not know how to incorporate it into their decision processes. Finally, evaluators must become aware of the vital role their personal style plays in evaluation utilization. Training procedures for evaluators must highlight the evaluator's role and increase interpersonal skills.

BIBLIOGRAPHY

Alkin, M.C. Improving Title I Technical Assistance: An Evaluation Utilization Perspective. Paper presented at the Annual Convention of the American Educational Research Association, Boston, April 1980.

Alkin, M.C., Daillak, R.H. and White, P. Using Evaluations: Does evaluation make a difference? Beverly Hills, CA: Sage Publications, 1979.

Anderson, J.I., Fishbein, R.L. and Stonehill, R.M. The Uses of Title I Evaluation Data in Local Decision-Making--A Federal Perspective. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 8-12, 1979

Apling, R.N. and Kennedy, M.M. Providing School Districts With Assistance to Improve the Usefulness of Tests and Evaluations. Cambridge, Massachusetts: The Huron Institute, September 1981.

Bessey, B.L. The Impact of the USOE Title I Evaluation and Reporting System: TAC Perspective. Paper presented at the annual convention of the American Educational Research Association, San Francisco, April 8-12, 1979.

Boruch, R.F. and Cordray, D.S. An Appraisal of Educational Program Evaluations: Federal, State and Local Agencies. Evanston, Illinois: Northwestern University, June 30, 1980.

Cichon, D.J. and Herr, G. Use of Title I Evaluation: Results in an Urban School District (Technical Investigation Final Report). Hampton, New Hampshire: Region I Title I Technical Assistance Center, RMC Research Corporation, July 1981.

Crowley, D. LEA Usage of Title I Evaluation Data for Decision Making and Program Planning. State of Indiana: Division of Compensatory Education, Department of Public Instruction, March 1981.

David, J.L. Local Uses of Title I Evaluation. Palo Alto, California: SRI International Educational Policy Research Center, July 1978.

Forgione, P.D. Jr., Kaplan, B.A., Orland, M.E. Evaluation of Compensatory Education Programs: Problems, Promising Strategies and Recent Trends. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 8-12, 1979.

Goettel, R.J., Kaplan, B.A., Orland, M.E., Forgione, P.D. Jr. and Huff, S.M. A Study of the Administration of the Elementary and Secondary Education Act (ESEA), Title I in Eight States: Volume 5 - A Synthesis Report. Syracuse, New York: Syracuse Research Corporation, October 1977.

Hamilton, G.E. Title I Evaluation From the Perspective of the Local School. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 9, 1980.

Hansen, J.B. and Oxford, R.L. Effects of the Title I Evaluation and Reporting System: TAC Perspective. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 8-12, 1979.

Kennedy, M., Neumann, W. and Apling, R. The Role of Evaluation and Testing Information in Title I Programs. Cambridge, Massachusetts: The Huron Institute, September 1980.

Kennedy, M., Apling, R. and Neumann, W. The Role of Evaluation and Test Information in Public Schools. Cambridge, Massachusetts: The Huron Institute, August 1980.

Millman, J., Paisley, W., Rogers, W.T., Sanders, J.R. and Womer, F.B. Performance Review of USOE's ESEA Title I Evaluation Technical Assistance Program. Washington, D.C.: HOPE Associates, 1979.

National Center for Educational Statistics. School District Needs for Title I Evaluation Technical Assistance. Washington, D.C.: U.S. Department of Health, Education, and Welfare, Educational Division Fast Response Survey System, April 1979.

Patton, M.Q. Utilization-focused evaluation. Beverly Hills, CA: Sage Publications, 1978.

Stecher, B.S., Alkin, M.C., and Flesher, G. Information Use in Elementary School Decision Making (CSE Report No. 160). Los Angeles: University of California at Los Angeles Center for the Study of Evaluation, December 1981.

Stonehill, R.M. and Groves, C.L. U.S. Department of Education Policies Toward ESEA, Title I Evaluation Utility: Changes in Attitudes, Changes in Platitudes.

Troy, J. Toward Greater Utility of Title I Evaluation: Technical Assistance Center Perspective. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 8-12, 1979.

APPENDIX A

Data Collection Instruments

District Name

INDIVIDUAL INTERVIEW FORM

Circle category of respondent:

Board Member/Superintendent or Asst. Superintendent/Title I Coordinator or Federal Program Coordinator/District Title I Evaluator/Principal/School Title I Coordinator/Teacher/Parent Member of School Site Council/Other_____

(Please try to do at least 5 interviews per district -- preferably 6. For all districts, please include in the interviews the three categories underlined above.)

Utilization Instances

(Use the matrix included on the District Summary Sheet to identify utilization occurrences -- #1-#28.)

What were the two most "clear cut" instances of evaluation utilization identified by this respondent?

Utilization Occurrence "A"

Indicate Categories (1-28):_____

Briefly Describe:

Relevant Quote:

How much time elapsed between presentation of evaluation data and the utilization occurrence?

Explain if required.

Utilization Occurrence "B"

Indicate Categories (1-28):_____

Briefly Describe:

Relevant Quote:

How much time elapsed between presentation of evaluation data and the utilization occurrence?

Explain if required.

Additional Utilization Occurrence

Was there an "unusual" instance of evaluation utilization that you might wish to report?

If so, indicate categories (1-28):_____

Briefly Describe:

Relevant Quote:

ADDITIONAL REMINDER: Find out about TAC impact.

District Name

DISTRICT SUMMARY SHEET

For the District: Please summarize the number of different utilization occurrences (situations where evaluation impacted) described by all of your interviewees. Place the total number in appropriate cells dependent upon the evaluation data source and the type of use.

Data Derived From

Data Used For	Needs Assessment	Title I Performance Report	Local Process Evaluation	TIERS	TESTING RE Student Selection	Other LEA Testing	Other (Specify) _____
Attitude Change	1	2	3	4	5	6	7
Situation Recognition	8	9	10	11	12	13	14
Situation Action: Adm./Org.	15	16	17	18	19	20	21
Situation Action: Curr./ Instr.	22	23	24	25	26	27	28

1) Identify by cell #, which occurrence was the most clear cut unequivocal example of utilization. Cell # _____.

Since this occurrence is probably described on a subsequent page, please identify by occurrence number (# _____) and interviewee category (_____).

If further description is necessary, please provide.

2) Identify by cell #, another clear cut occurrence of utilization. Cell # _____.

Please identify by occurrence number (# _____) and interviewee category (_____).

If further description is necessary, please provide.

3) Was there another identifiable occurrence of utilization that was more complex (e.g., used data from multiple types of data sources)? If so, identify the occurrence by number (# _____) and interviewee category (_____).

If further description is necessary, please provide.

4) What was tne most unusual occurrence of utilization? (Occurrence # _____.) (Interviewee category _____.) Indicate why it was so unique. If further description is necessary, please provide.

5) In the instances of impact that you have found, what are the major factors associated with evaluation utilization?

(Do not feel constrained by particular research frameworks or points of view -- quite simply, what factors did you find?)

We turn now to the question of the impact of Technical Assistance Centers (TACs). Please provide the follcwing information:

6) Did the district respondents indicate that they received advice/ assistance from the TAC Centers? Yes ___ No ___.

If so, which respondents so indicated?
Circle: Supt./ Dist. Title I Dir./ Dist. Evaluator/ Principal/ School Title I Coord./ Teacher

7) What was the nature of that assistance?

- Workshop Training? _____

Specify content

How did it help to change and/or improve the program?

- Technical Assistance: Evaluation Procedures? _____

Specify

How did it help to change and/or improve the program?

- Technical Assistance: Test Interpretation? _____

Specify

How did it help to change and/or improve the program?

- Technical Assistance: Other? _____

Specify

How did it help to change and/or improve the program?

APPENDIX B

Middle City, Midwest: A Study of Title I Evaluation Use

Larry Braskamp*

University of Illinois

January 1982

*with the assistance of Jack McCurdy

The Setting

Middle City is the hub of one of the richest farm areas in the agricultural belt of the Midwest, perching like a citadel amidst the corn and grain fields that reach as far as the eye can see in all directions. There are no hills or mountains here, no natural barriers to interrupt a gaze across the flat, endless expanse of green stalks that shimmer in the hot winds rolling across the plains in the summertime. The only embellishments are the tall office buildings and bulky factories of Middle City where business and agriculture join forces to produce the dominant farm economy of this Midwestern state.

Most of the land area within Middle City's boundaries is taken up by the apartments and homes that adjoin downtown, the residential suburbs that have spread outwardly and the industrial area to the north where factories turn out much of the farm equipment that supplies much of the Midwest. This geographical layout of the community has produced distinctively different living patterns and has given Middle City two faces instead of one. The middle class and more well-to-do live in the relatively quiet and often fashionable suburbs while the poorer inhabitants are clustered around the industrial section where housing is often shabby and the crime rate is high.

The School District

Middle City's school system, which has nearly 30,000 students and 49 schools, also reflects these divisions. Most of the district's low-income and minority students are concentrated in the schools in the northern industrial section. In all, about 30% of the enrollment is classified as low-income, and the minority students comprise about 15% of the total enrollment, including about 13% black, 2% Latino and 1% Asian-American. Minority concentrations range from 2% to 93% in the individual schools with nearly all the racially impacted schools located in the northern area.

Since the 1960's, the schools in this area have been served by a Title I program, which is the main focus of this evaluation study. In Middle City, the Title I program has four components: elementary reading, elementary mathematics, secondary reading and mathematics and early learning centers. During 1980-81, the combined components served some 3,200 of the estimated 6,000 students eligible for Title I services in 29 schools at a cost of $1.8 million.

Nearly 2,000 elementary pupils were in the reading component staffed by 55 teachers plus aides, and about 1,300 students received math instruction from 35 teachers plus aides. The secondary program included about 475 students and 13 teachers. About 480 four-year-olds were enrolled in the early learning center program, which was staffed by five pre-school

teachers plus a psychologist, nurse, speech clinician and aides on a part-time basis. The Title I program staff included a supervisor and four coordinators, each responsible for overseeing one of the four components.

Student eligibility for Title I is determined through the use of standardized tests in reading and mathematics which measure the youngsters' ability to attain a series of prescribed performance objectives. For instruction, Title I classes use a management system that covers both reading and mathematics. It includes performance objectives and criterion-referenced tests that enable teachers to monitor student progress toward mastery of materials. Norm-referenced tests also are administered to Title I students in the fall and spring to measure achievement and gauge program effectiveness.

The Evaluation Department

The district's evaluation department was established about 10 years ago and was purposely assigned to the administrative services division--outside the instructional division--to assure its independence in evaluating educational programs. Jack Brown, its director, has headed the department since its inception. He is a long-time district employee and former principal with no formal training in evaluation. Borwn's staff includes three evaluators, one a former administrator and two with evaluation backgrounds in psychometrics.

One of the department's main responsibilities is to write annual evaluation reports on each of the four Title I components. Each 20- to 30-page report includes a description of the component, selection criteria and the instructional program. But the most important part of the report is a discussion of the component's objectives. This is divided into two sections, one covering process objectives such as administration of the component and services provided. The other deals with product objectives--student academic performance and learning progress. This section evaluates the degree to which objectives adopted by teachers in each component have been achieved and is based largely on test data. The report concludes with a narrative that summarizes the overall performance of the component and carries three recommendations for improvement.

In addition, the evaluators make on-site evaluations and try to visit each Title I teacher during the year. They observe the programs and check documentation for attainment of component objectives. Based upon these visits, they write memoranda evaluating the programs which are submitted to the Title I supervisor. These memos are separate from the evaluation department's reports and are more subjective. They are turned over to the schools by the Title I supervisor for their study.

The evaluation department also supervises the administration

of the norm-referenced tests in the spring and fall, but Title I personnel administer and score tests used for determining eligibility of students. As a matter of practice, the evaluators also work closely with Mrs. Kohl, the Title I supervisor, the Title I four component coordinators, and teachers on a wide range of matters affecting the program.

Evolution of Evaluation in the District

During the past few years, the evaluation department has played an increasingly active role in the district. Evaluators initially spent much of their time helping Title I administrators and teachers write process and product objectives. In some cases, they would simply react to the objectives written by Title I staff and suggest changes. Each spring, evaluators also would assist Title I committees composed of teachers and administrators in the preparation of the next year's funding proposals. But, with few exceptions, the evaluation department did not involve itself in assessing the attainment of the objectives and the performance of the program.

Since 1978, however, the evaluation department has gained in stature. That was when the superintendent added two persons to the department and announced his commitment to a new elevated status that evaluation would be given in the district. He did so following expression of public concern about the success of academic programs in the district.

Then, in 1981, a new superintendent was hired, adding new support to the evaluation function. He publicly advocated the use of evaluation in decisionmaking at all levels. His focus is on the accountability of programs and the use of evaluation tools to assess program performance. He requested the evaluation department to provide him with recommendations on the value of programs and even to suggest whether they should be continued. As one school board member put it, "The superintendent made evaluation a top priority. It is working out quite well."

The new administration's perspective on evaluation was summed up by an assistant superintendent, Dr. Blackwell:

> We want to get a well-oiled process for helping students achieve objectives. Evaluation makes this process more definite. It makes us more sensitive to specifics in management. No assumptions are made that one does a good job. We do it by the pre- and post-test differences. Evaluation does not assume. It shows evidence. Through evaluation we are trying to make him (the student) measurably better.

One evaluator underscored the difference. "Now we are able to do a careful monitoring of programs. The new superintendent knows the potential benefits of an evaluation and his more business-oriented approach to organization gives our department a more prominent role." Another evaluator commented: "In the last three to four years, the emphasis has changed to finding out more about what is going on in the programs. We have recently done less monitoring of the process and more on performance. To the teachers, we are viewed as a source of information."

At first, some teachers were wary of the administration's concept of using evaluation as "an important management tool." They were defensive about the evaluators moving into program evaluation and there was resistance. When the Title I staff realized that evaluation could be valuable in helping them improve their program, however, their attitude began to shift. As the Title I supervisor said, "The evaluation department is here to stay, and the staff and teachers know this."

Dr. Blackwell said that although the evaluation department may never be loved, it performs a job that must be done. The evaluation director "knows he will be soundly resented," the assistant superintendent said. "He is brutally honest. He wants the truth to be known. He tries to give preliminary information so there is an opportunity to get things corrected. He would like to give favorable evaluations if he can, but he wants the kids to have a fair shake. This way everyone wins in the long run."

There is general agreement in the district that evaluation has had a significant impact since its role was elevated several years ago. It has led to improvements in the areas of instruction, program content and in information exchange in the system.

Utilization

Impact on Instruction

One of the first things that showed teachers the value of the evaluation department was its help in organizing reading test questions and answers so they were of use in planning instruction. The test scores had long been routinely distributed to teachers but that format made it virtually impossible for teachers to relate the results to strengths and weaknesses of students in the various skill areas. As Mrs. Kohl, the Title I supervisor,explained, the total scores "may be good for the newspaper, but there is not enough information for us to help improve the program."

So the test questions and answers were organized into clusters of reading skills. These included vocabulary, literal specific, literal global, inferential specific and others. In this way, teachers were able to identify how each student performed in the skill areas and to plan individual instruction accordingly. "The scores in clusters," the Title I supervisor said, "are much more useful to us and to the teachers because they give us information about the problems and how we can focus on these problems such as designing inservice training programs for tackling them." The elementary reading coordinator said the cluster scores "have helped the teachers' attitudes toward evaluation as well." A Title I teacher agreed. "The clusters are a good

idea," she said. "They give me an idea of where I should work."

But even more important has been evaluation's influence on the instructional approach followed in the district. The evaluation department recommended adoption of a mastery approach which was first used in the Title I program. It is now in the process of being implemented throughout the system. It is known as the Performance Improvement Program (PIP), which consists of prescribed learning objectives that students are expected to master in sequential order. It also includes criterion-referenced tests to enable teachers to monitor student progress and focus instruction on detected deficiencies. PIP is termed an instructional management system and is designed to assure mastery of learning objectives and reduce student failure. Thus, the evaluation department's recommendation is expected to produce a fundamental reorientation of instructional strategy in the district.

In another case, an analysis of test results by the evaluation department has led to changes in some of the instructional methods used in the Title I program. Its origin lay in the department submitting its report on reading and math scores from the norm-referenced tests to the Title I supervisor, Mrs. Kohl. The data showed that while other grades had met their test score objectives, the fourth grade had not. Mrs. Kohl, and her staff went to work to find out why. They talked to fourth grade teachers and observed their classes. They discovered that reading was only covered once a day in

most fourth grade classes compared to two to three times a day in the earlier grades. The scores of the fourth grade pupils also were examined by clusters, and it showed that the fourth grade was much lower in the area of comprehension skills. The Title I team found that while the fourth grade was expected to begin placing greater emphasis on reading comprehension, the teachers felt it was a difficult skill to teach and they lacked confidence in their own abilities to do it well. As an outgrowth of their investigation, in-service sessions were organized to demonstrate to teachers ways of teaching reading comprehension, including exercises in writing stories and methods of testing students about story themes.

Impact on Program Content

The most substantive effect of evaluation on the Title I program concerned a decision to cut reading labs based on use of data to evaluate their effectiveness. As part of the Title I reading program, students in some schools were assigned to reading labs for extended periods of time to enable teachers to provide intensive remediation on an individualized basis. Actually, the lab time devoted to reading was considerably more than was required by Title I guidelines, and the labs were more expensive than other facets of the Title I program on a per-pupil expenditure basis. Using data provided by the evaluation department, the Title I staff and teachers discovered that test scores of youngsters assigned to the read-

ing labs had not improved significantly over the previous four years of experience with the labs. Based on this information, it was decided to terminate the labs. Teachers and aides who staffed the labs were shifted to other tasks, and funds for the labs were reallocated within the Title I program. In the place of the labs, the schools began pulling out students from regular classes for special reading instruction. The Title I staff and teachers chose pull-out instruction as a substitute for the labs because other Title I schools had reported successes with it.

In another instance of program impact, test results from the evaluation department produced a shift in assignments of pupils in the Title I program. Scores from the norm-referenced tests which has been submitted to the Title I program coordinators showed that in the fourth grade, students new to the program were making markedly greater gains in reading than other students who had been in the program longer. Title I coordinators and reading specialists then made observations of the fourth grade classes. They determined that many of the teachers were devoting more of their attention to the newer students. "They were not challenging all the students to the best of their abilities," Mrs. Kohl said. "The teachers were not helping students (who had been in the program before) as much as they should have." Subsequently, a number of the students who had been in Title I for some time

were removed and placed in regular classrooms. At the same time, in-service sessions were organized to show teachers how to teach on a more individualized basis and to equalize their time and efforts among all students.

Over the years, the evaluation department has also had a crucial influence on Title I in terms of the quality of the program objectives that it establishes for itself. One of the evaluators' main roles has been to critique the objectives written by teachers and Title I staff, both in the process and product areas. The evaluators have been instrumental in the development of objectives that are much more defined, measurable and clearly-stated. As a result, these accomplishments have tended to produce a Title I program that is more structured and uniform across grades and among services. Although teachers still have the flexibility of using the teaching techniques they prefer, the Title I supervisor said a more structured program based on clear objectives now provides greater assurance that children will achieve more satisfactorily. "It has made us much more accountable," Mrs. Kohl said.

Impact on Information Exchange

The evaluation department has been responsible for development of an information system for the district, which provides valuable data for decisionmaking. Evaluators designed a computer-based system which enables teachers and staff

to monitor student progress in various academic areas. For the Title I program, it provides more complete records of the activities of Title I students and how they spend their instructional time. For example, information is compiled on when and for how long students have been pulled out of regular classes and if this affects overall learning in various subject areas.

The plan that is used to disseminate the Title I evaluation reports also has produced a communication network that has involved all staff from Title I teachers to top administrators in the evaluation process. It begins with the annual evaluation reports being prepared by the evaluation department. To recap, these reports focus on the extent to which objectives of the Title I program have been met and include recommendations for improvement from the evaluators. The reports then are submitted to the Title I office, which sends them through the component coordinators to the schools. At the schools, teachers are asked to respond in writing with their reactions and suggestions. Those responses along with the reports then go back to the coordinators, who also write their comments. Both sets of reactions are incorporated into an overall response from the Title I supervisor, who submits the package to the assistant superintendent of instruction and the superintendent. It is discussed at meetings of the district's top administrative staff, and a summary version is

transmitted to the school board.

The process of circulating the reports is considered valuable in several ways. It encourages, if not requires, key staff members--primarily teachers--to read and comment on the evaluations. In many instances, it leads to group discussions of the data and their implications, and it generates recommendations from various perspectives along the network. Top administrators are afforded comments from evaluators as well as teachers and Title I staff in a healthy give-and-take over program performance. Multiple suggestions for improvements also come out of the exercise. One evaluator said the chief benefit has been the active involvement of teachers in evaluating the program. "In the past," the evaluator said, "the teachers responded to our recommendations rather passively. But now they present their side of the story." This communication process has led to numerous refinements and adjustments in the Title I program. But more importantly, it has made many people close to the Title I program an integral part of evaluating its performance and taking responsibility for thinking about ways to improve its performance.

Observations

The reasons for the uses of evaluative information in Middle City, Midwest relate to and flow from the district's organizational culture. The valued norms and expectations of

the high level administrators, language used by staff, teachers, and administrators in talking to each other about instructional concerns, the organizational structure of Title I, and most importantly the type of instructional delivery system, all reflect a commitment to an objective-based approach to instruction. Thus test data are valued; they are build into the selection of Title I students, used to monitor student progress, used prescriptively by teachers through the use of clusters of items classification scheme, and used as clues to further investigate problems implied by the test data. In sum, the people with formal power have made testing a part of the modus operandus for instructing students.

The evaluative information is predominantly the test data measuring student achievement both through the use of the criterion referenced tests tied directly to the delivery of instruction and the scores (pre-post) of the norm referenced standardized tests. These data provide the formal information base for the district. The data are trusted, and the role of the evaluators is one of collecting and reporting this type of information. The evaluators are viewed as information providers--as technical staff. Thus they are relatively free from criticism. The test data, since they focus on student achievement, keep discussions about effectiveness at the programmatic and curricular level. Thus the threat of evaluators has been minimized and "evaluation as a management tool"

is the perception espoused by the potential users. And its actual use seems to verify the validity of this oft used phrase.

The value of the Evaluation Department varies depending upon the criterion of usefulness. The evaluators' usefulness is perhaps best viewed in terms of making a contribution to the evolution of the district-wide objective based curriculum, the Pupil Improvement Performance. The Evaluation Department's impact is indirect and not always linear, but if one takes a long term view the influence is substantial. The evaluators' role is one of offering suggestions for integrating testing in the delivery system, providing techincal expertise and remaining "objective" data providers. Their influence is largely due to the support they receive by their intended audiences, particularly those in power.

The use of evaluative information is dependent upon the purpose for which it is to be used. Standardized test data were used in three ways. First, teachers used it to learn more about their students. This use by Title I staff and teachers was enhanced by the active involvement of the Title I staff in establishing the clusters of items. Thus partial ownership of the scheme facilitated acceptance of this portrayal of the data. Second, standardized test data were used in a direct allocative way, although this use was not frequent. Its use however was precipitated by the need to reduce the budget.

The best example is the elimination of the reading labs. Test scores and Title I teacher comments were given as the sources of data used in this decision. Thirdly and most frequently, standardized test data triggered further investigation of a problem identified by the test data. Naturalistic investigations were done to more fully understand the problem identified. Here informal and qualitatively oriented methods of inquiry (e.g., observation of classrooms, interviews with teachers)were conducted. The use made of these ad hoc evaluations was often action. The reason for high use was perhaps due to the fact that these evaluations were commissioned by the potential users and often conducted by the Title I staff. Thus the information need for evaluations of the identified problems was extremely high.

In sum, evaluation in the Middle City school district has affected its Title I program in numerous direct and subtle ways. The orientation of evaluation toward instructional activities has generated a more critical and demanding view of the program from teachers as well as administrators. It has produced some concrete instances of structural and programmatic change that resulted in more effective use of resources and greater likelihood of learning progress among children. But most of all, it has given the school district and its staff a feeling that there are systematic ways to examine a program, primarily through the use of instructional objectives based test data.

APPENDIX C

Big City, East: A Study of Title I Evaluation Use

Richard Daillak

St. John's University

January, 1982

The Setting

Context

Big City is marked by striking contrasts. From its concrete, glass, and steel office towers, there unfolds a panorama of affluence juxtaposed against poverty, power against powerlessness, chic brownstones against scarred tenements. And, too, there is that vast middle ground of average citizens living their daily routines in a way that city dwellers almost anywhere would recognize.

To accommodate the sheer size and diversity of Big City, the schools decentralized almost a decade ago, forming smaller self-governing "regions." The central board and administration retained control of the regions' overall budgets, however. In addition, secondary education throughout the city is controlled from the central office.

Decentralization was often painfully marked by contention for power among parents, school administrators, and teachers. Region D, the site of this case study, emerged relatively unscathed from decentralization. As the teachers' union liaison to the Region's headquarters put it, "Region D never went off the deep end, the way some did," and the community did not succumb to the "let's take over the schools" spirit that had struck elsewhere. From the point of view of teachers and administrators, Region D was fortunate to have escaped the turmoil and instability which, they argue, detract

from productive educational work.

Still, Region D is not without challenges. It serves several thousand students of quite diverse backgrounds. Within its boundaries are posh residential neighborhoods and squalid housing projects. It can boast of one of the 10 best schools in Big City; it also operates a school with one of the highest percentages of Hispanic students. Overall, the Region serves a student population that is approximately 60 percent Spanish surname and 20 percent black.

Region D receives several million dollars in state and federal funding for a variety of special programs. Besides bilingual education, the Region operates basic-skills oriented compensatory education programs, which are funded through Title I of the Elementary and Secondary Education Act (ESEA) and through state aid. In the early grades (1 and 2), the "Early Childhood" compensatory program includes pull-out and laboratory service for small groups of children needing reading and math assistance. The precise organization of this program varies from school to school, and some administrators and teachers feel that this program needs strengthening.

The upper elementary grades (3 through 6), on the other hand, have been operating a highly structured program in reading and math: the HILS (High Intensity Learning System) program, from Random House. Satisfaction with HILS has

been strong, so much so that the program has been identified as a model for other regions. HILS math appears particularly successful--a mixed blessing since math scores have been raised to the district-wide average reducing the apparent need for compensatory assistance in mathematics. Faced with the imperative to make budget cuts, Region D has chosen to discontinue HILS math and concentrate funding on the reading component alone.

In the Region's junior high schools, reading instruction has been a major compensatory focus, through the vehicle of the "Corrective Reading" program, which serves students individually or in very small groups. Recent evaluations of this program have been positive, although areas for improvement have been cited.

Besides the major programs just mentioned, Region D's schools support quite a number of less widespread special activities. There are laboratories and skills centers of various types sprinkled throughout the Region. Some are the remnants of previously widespread efforts; others trace their origins to the particular enthusiasms of local teachers or principals, or to the unique needs of a school's student population.

The Region D Evaluator

The Director of Funded Programs Evaluation in Region D, Harold Green, has held his position since 1974. Now in his

forties, he is, in years on the job, the dean of region evaluators in Big City. But, for Mr. Green, evaluation was a second career; originally, he was in one of the helping professions. In the late 1960s, he sought a change of career and completed a Masters program in educational research and evaluation at a local university. Mr. Green worked for two years for the university's research bureau, conducting evaluations for Region A in the city. Then he spent a year working directly for Region A as their in-house evaluator before moving to his current position at Region D.

Mr. Green's university training prepared him for the technical demands of his job. But his personal characteristics--intelligence and a streak of good-natured independence and assertiveness--truly suited him to it. More than that, Mr. Green created an evaluation role in Region D that fit his own personal style.

Evaluation in Region D

Prior to 1974, Region D relied upon external consultants for evaluation services. The Region worked with consultants from two local universities, as well as from educational institutes in the Big City area, but was less than fully satisfied with these consulting arrangements. According to the Region Superintendent, an evaluator on site could more easily provide them with timely assistance, for example, when

they were preparing program proposals or when a question of topical concern arose.

Mr. Green undoubtedly helped persuade the Region's administrators of the value of in-house evaluation, for he approached Region D before it had been decided even to create such a position. Mr. Green read an employment advertisement for a new Director of Funded Programs in Region D, called the Region office to ask if they had considered establishing an evaluation position, and was, in fact, hired before the Funded Programs directorship was filled.

Mr. Green secured assurances from the Superintendent that his evaluation work would be independent of any pressure, either from within the Region or from citywide administration. He established a hiring policy that no one who lived within the Region could work on his staff in an effort to guard his independence. And when he sought editorial comments on his evaluation reports, he did so through a formal exchange of memoranda. Mr. Green states that on his first full evaluation report (for 1974-75), 30 changes were recommended by the Superintendent, of which Mr. Green made seven.

Mr. Green was not, however, insensitive to the interpersonal or political side of evaluation. Upon arriving in Region D, he deferred any work in the schools for six weeks, until a meeting could be arranged with the union stewards for all the Region's schools. At that meeting, he explained

his "ground rules" for evaluation--anonymity for individual teachers and schools, and independence from building principals--and secured the cooperation of the union. Mr. Green also has given individual schools the opportunity to comment on negative evaluations prior to submitting his reports. He says he hopes that teachers and administrators will be able to acknowledge his fairness, even if they do not always agree with his conclusions.

A hallmark of Mr. Green's approach to evaluation in Region D has been his regular use of subject-matter experts as consultants. Beginning in 1974-75 and continuing through 1978-79, Region D's evaluation reports incorporated both test score results--the standard evaluation data required by funders--and narrative evaluation critiques based upon on-site observations by these consultants. A 25 percent budget cut forced the suspension of the site observations, and the 1979-80 evaluation consisted solely of the TIERS (Title I Evaluation Reporting System) and other achievement test data required by external agencies. (The 1980-81 evaluation is described in the section to follow.)

Title I evaluation in Region D follows TIERS Model A: students are pretested (in early fall) and posttested (in late spring), and normal curve equivalent (NCE) scores are compared from the two testings. Selection of students for Title I is based upon the preceding spring's test results,

among other criteria, thus helping compensate for the regression effect. Through 1981, the California Achievement Test (CAT) was used for English-language testing. Mean NCE scores on the subtests (Vocabulary, Reading Comprehension, Math Computation, Math Application) were reported by program and grade level within program. T-tests were employed to assess the mean differences between the two testings at each grade level, and statistically significant differences were noted in a brief narrative summary. Statistical tests used individual student scores as the unit of analysis. Some additional descriptive summaries were prepared. The test score analyses were prepared by the evaluator, Mr. Green.

Narrative critiques of specific programs (Early Childhood, HILS, etc.) were prepared by the subject-matter consultants and edited by Mr. Green for inclusion in the final evaluation report. These critiques described program strengths and weaknesses and gave recommendations for program changes. Quantitative analyses of student performance which would rigorously support the consultant's recommendations were not routinely conducted, however.

Drafts of the evaluation reports were presented to the Superintendent, for written comment. Copies of the finished reports were submitted to the Superintendent, the Region School Board, the Region's Director of Funded Programs, the Region Parent Advisory Council, Region Office coordinators

for reading and mathematics, and the Region's schools. Mr. Green usually discussed the evaluation results informally with the Superintendent, the Funded Programs Director, and the coordinators, in addition to submitting the written report. Generally, Mr. Green orally presents the results to the Parent Advisory Council, as well.

Besides receiving Mr. Green's evaluation, the compensatory programs have also been scrutinized by the Parent Advisory Council. Parents visit an average of one school per month. Typically, the parents observe teachers' work, examine the material in use in classes, and discuss the program with staff.

Over the years, Mr. Green also has provided special services, on an as needed or as requested basis. He has prepared evaluation plans for program proposals, and he has provided quick analyses of issues raised by Region administrators. For example, he was asked to investigate a test score decline a few years ago, and his analysis attributed the decline to certain sixth grade classes (as discussed below, in the section on Utilization). His special analyses are issued as memoranda or small reports.

The 1980-81 Evaluation

In 1980-81, certain changes were instituted in Region D's evaluation. Most significantly, a new form of narrative critique of the compensatory programs was developed. Rather

than sending consultant-observers into schools and classrooms--a method too costly to reinstitute--Mr. Green brought classroom personnel to the consultant, by means of interviews at the Region Office. Group, and a few individual, interviews were conducted with HILS and Early Childhood program teachers, TESL (Teaching English as a Second Language)/Bilingual Teachers, teachers in the state compensatory education program, members of the parent advisory council, Region coordinators for the special programs, principals of the Region's schools, and the Region's Superintendent. The participants were asked to describe the strengths and weaknesses they saw in the programs, to recommend changes and model practices, and to assess, as candidly as possible, the real effects wrought by the programs and the losses, if any, that would occur if their programs were discontinued. The results of these interviews were summarized by the consultant (who had served as an observer in previous years) and incorporated as the initial 10 pages of the evaluation report.

Less obvious from the written report was a change in the way that test score data was prepared for submittal to the state. Previously, teachers in the schools had had to complete individual student data reporting cards based upon the test score results they received. For 1980-81, Mr. Green contracted out the reporting process, significantly reducing the amount of time that teachers had to devote to record

processing, and producing a net saving to the Region of a few thousand dollars.

As before, the evaluation presented test score results by program and grade level. It also included, as a separate section, a consultant-produced evaluation of one school singled out for special attention because it had recently implemented a "school-wide" Title I program. The special evaluation, promised in the program application, reported NCE and raw score gains by grade level within the school involved, summarized classroom observations and staff interviews, and recommended program continuance.

In 1980-81, Mr. Green conducted two special analyses apart from the regular evaluation. The first of these was carried out at the Region Superintendent's request and involved a review of student performance data from the last several years. To prepare for forthcoming hearings on Title I services, the Big City Schools' headquarters had asked the Region D Superintendent for data that might be used to demonstrate Title I effectiveness. The Superintendent passed the task to Harold Green, who examined student achievement data over the last decade. In a memo to the Superintendent, Mr. Green presented several indicators of an upward trend in student performance, a trend which he argued should be attributed in large part to the presence of Title I and other special aid. In 1972, he noted, only about 20 percent of

the Region's students had been reading at or above grade level. This percentage had increased to about 45 percent by 1981. Scores from Big City as a whole had moved upward in this same period (from a little more than 30 percent, to 50 percent at or above grade level), but Region D's improvement was even steeper. In mathematics achievement, Region D had essentially closed the gap between its performance and citywide averages during the same period. Overall, Region D had moved up in the relative rankings of Big City's several Regions, and an examination of the rankings over the last several years showed that Regions with large Title I programs had shown the greatest improvements in relative standing.

Another analysis was initiated by Mr. Green himself. Student achievement data was collated with information on the average number of times per week that each child was served by Title I. After adjusting for pretest performance (using ANCOVA), "optimum" service frequencies were determined for each grade level. These were five times per week in Grades 1 and 2, four times per week in Grades 3 through 6, and seven times per week in Grades 7 and 8. A tentative rationale for these findings was presented, and the optimum frequencies were recommended for implementation in subsequent years.

1981-82 and Future Evaluations

The group interview strategy was considered very successful

and is being continued for 1981-82. A change in evaluation tests has been instituted, however. Beginning with the fall 1981 pretesting, Region D will be using the CTBS (Comprehensive Test of Basic Skills) for Title I evaluation. The evaluator, Harold Green, contended that teachers were "becoming tired" of the previous test, the CAT. Some may also have begun to teach to the test.

Utilization

Special Analyses

Harold Green's special analyses apparently influenced the local audience. The recent review of student achievement trends, conducted at the Region Superintendent's and Big City Schools' request, was presented by a top Big City Schools' administrator at a state legislative hearing on compensatory education programs. Mr. Green's self-initiated analysis of service frequency was discussed with the Region Superintendent and the Director of Funded Programs. They are considering implementing a policy calling for service at a frequency of at least five or six times per week in grades seven and eight, and they may present the findings to principals in order to encourage compliance.

As mentioned previously, the Region Superintendent asked Mr. Green, a few years ago, to analyze a sudden test score decline affecting the Region and the entire City. The analysis

revealed that in Region D only the 6th grade had declined, and, then, only those 6th grades within grades 6 - 8 middle schools. Subsequently, the Superintendent reassigned 6th grade classes to K - 6 elementary schools, rather than middle schools, in all but a few cases within the Region.

Narrative Critiques

Whether derived from observations or the more recent group interviews, the consultant-generated narrative critiques of Region D's special programs appear to have gained a sympathetic hearing. For example, the following instances of utilization were mentioned by one or more respondents in the Region.

- In the group interviews, teachers indicated that they thought the math centers established in the HILS math program would be useful for regular teachers even if the HILS math tutoring was discontinued. This information was one consideration in the ultimate decision to stop HILS math.
- Interaction among TESL/Bilingual teachers during their group interview was described (by the Funded Programs Director) as especially productive. TESL was a loosely structured program in Region D, and the group interview provided an uncommon opportunity for teachers to share ideas, problem solution, concerns, etc.
- A consultant-observer identified two distinct teacher styles ("managerial" versus "circulating") in HILS reading classes, criticized the managerial style, and recommended the circulating style as more effective. Subsequently, the Region held workshops to encourage HILS teachers to adopt the circulating style and spend less time at their desks.

- Another consultant observed math labs and recommended greater use of manipulatives. Workshops were held to promote this recommendation.

- A consultant observed learning disabilities programs within two schools, one employing a self-contained classroom, the other, a resource room. The self-contained classroom was strongly criticized. The principal at the criticized school read the consultant's report, declined Mr. Green's offer of a second opinion, and changed to a resource room approach.

- A consultant observed classes in one of the Region's better schools and complimented a teacher on her excellent work. The school's principal reported that this was important to teacher morale--especially because it came from an impartial expert.

In general, the narrative critiques seemed to be read carefully by the subject-matter coordinators in the Region D Office. Acceptance at the school level was more difficult to assess independently in this brief study; Region Office respondents asserted that the consultants' suggestions usually were well received by teachers, in part because teachers perceived the consultants to be genuinely knowledgeable.

Standard Test Score Data

Local utilization of the externally required achievement test data was less extensive. The test data were used for occasional special studies, as has been described in a preceding section, but routine use of the evaluation reports' achievement summaries (mean NCE scores by program and grade level) was not apparent. The evaluator, Mr. Green, stated

that the TIERS data itself was not particularly useful, although he had recently seen some analyses, generated by the state department of education, which he felt might be more interesting. He did not have copies of these new analyses, however.

The Director of Funded Programs said that individual building averages were considered during program planning and resource allocation. (Building mean scores were not reported publicly.) Schools with the lowest mean test results often received greater assistance than those schools with higher performance. Planning was asserted to be comprehensive, taking into consideration the full range of special resources (ESEA Title I, Title VII, state compensatory and bilingual aid, etc.) available to a school.

In the past, it has been particularly difficult to distinguish the influence of Title I testing from the effects of district initiated achievement tests. Region D's Title I evaluation has employed the CAT--the test selected by the district, as well--and the district's spring testing has served as the Title I posttest. The change to CTBS for Region D's Title I data, described in an earlier section, will alter this situation.

Evaluation's Informal Influence

Region D's Director of Funded Programs said that a significant amount of evaluation's influence occurred informally,

in impromptu meetings or conversations he or the Superintendent might have with the evaluator, Mr. Green, throughout the school year. The offices of the evaluator, the Funded Programs Director, and the subject-matter coordinators were conveniently close to one another in the Region headquarters--and none were far from the Region Superintendent's office suite. Relations among all these personnel seemed open and friendly. This informal influence of evaluation, or of the evaluator, personally, could not be estimated well in this study, however.

Some Observations

The Success of This Evaluator's Approach

Harold Green appeared to have carved a niche for evaluation in Region D. Several elements of his evaluation appraoch probably contributed to this success.

Use of Consultants to Produce Narrative Reviews. Mr. Green did not feel that he had to do all parts of the evaluation himself. He said that he preferred consultants to full-time staff and always tried to select consultants with real expertise in the instructional tasks they were to critique. His emphasis on subject-matter experts, rather than evaluation methodologists, seemed well-received in Region D. The consultants' narrative critiques appeared to be more influential than the test score summaries in the evaluation

reports. The narratives made sense to administrators and teachers, and seemed to carry weight.

Cost-Consciousness and Responsiveness. Although Mr. Green enjoyed his independence, he was not unresponsive to Region circumstances or administrative requests. He was proud of his cost-consciousness, an attribute of considerable importance in recent years. When a 25 percent budget cut had been necessary, he had quickly accommodated by halting the school observations. His new group-interview strategy had returned much of the benefit at a much reduced cost. He had negotiated a contract for TIERS data processing at more than 25 percent under the initially budgeted expense and returned the savings to the general Title I budget. His preference for consultants was balanced by careful attention to office costs. His own office staff was modest: he had one paraprofessional assistant.

Mr. Green did his best to respond promptly and usefully to administrative requests for information, for example, in the special studies he did at the Superintendent's request. He worked closely with Region Office personnel, such as the Funded Programs Director and reading and math coordinators, and said that he thought it was a serious mistake for an evaluator to try to influence teachers directly. He was responsive to teacher and school sensitivities, as evidenced by his respectful approach to the teachers' union and his statement that he was concerned with evaluating "programs, not teachers or schools."

An Emphasis on Pithy Conclusions. Mr. Green described his job as that of "transforming mountains of data into useful information," and he valued concise interpretations and reports over thick research studies. The consultants' program critiques fit this style well, and Mr. Green's own analyses tended to be brief and to the point. It seems likely that this style won listeners among the personnel at the Region Office and contributed to evaluation information use.

The Question of Rigor

There is a degree of tension among the goals of conclusiveness, brevity, and rigor. The evidential base underlying the evaluation conclusions described in earlier sections could be criticized. The consultants' recommendations, for example, although informed by expert knowledge and first-hand observation, were not buttressed by hard evidence of the effectiveness or ineffectiveness of specific program practices. In the evaluator's special analyses, causal attributions beyond the actual data appear to have been made. For example, without experimental study one cannot confidently infer that children should be served some specific number of times per week or that 6th grade classes should be placed in elementary, rather than middle, schools. The situation, then, is one of evaluation conclusions grounded--imperfectly--in empirical data.

The Study in Perspective

The preceding review has focused tightly on evaluation and the evaluator in Region D. In the process, it is easy to overestimate the importance of evaluation. The case study telescopes seven years of evaluation work into one short account, concentrating on the occasions when evaluation had a significant, memorable, effect upon decision-making. What is missing is a full perspective on the programs involved and a better understanding of the relative contribution evaluation has made. Such an analysis is beyond the scope of this case study, of course.

APPENDIX D

Small Town, South: A Study of Title I Evaluation Use

Jean A. King

Tulane University

January, 1982

The effective use of Title I evaluation information is not limited to large urban school systems. The Title I Program in Small Town, South, integrates test results and other information into the daily functioning of its schools. How this happens is a tribute to the skill and hard work of two women and a school community that truly cares about its students.

The Setting

The characteristics of Small Town that are at once its strengths and potential weaknesses are readily apparent in a visit to the city. A drive to Small Town through the rolling southern prairie emphasizes its rural locale and the agriculture that is an important livelihood for many of its citizens. The city itself is small -- population near 8000 -- and, though it has no motels, it does boast four restaurants, an 1893 country courthouse "ugly enough to be beautiful", and more than 15 churches of various denominations. Thirty miles from the state capital and 45 miles from a big city, Small Town's location makes it an appealing home for many commuters.

The consequences of being a bed-room community are serious. Despite a county oil boom, the city has a low tax base. Its two largest employers are the school district and the city government, both of which are feeling the effects of federal cutbacks. The results of a recent census placed 28% of the county's family incomes below the poverty level, and because no major industries are creating large numbers of jobs in the city, this picture may not soon change. But, as

a Chamber of Commerce pamphlet exclaims, "Small Town, South, Looks Ahead!" The town has both a long history and high hopes for its future.

Some of these hopes rest with the Small Town Independent School District, where the proportion of 24% white, 65% Hispanic, and 11% black students roughly parallels the racial and ethnic mix of the town itself. In addition to a central administration building, the district maintains five schools for its 2600 students: a kindergarten; a primary school (grades 1-3); an intermediate school (grades 4-6); a junior high school (grades 7-8); and a high school (grades 9-12). The district's newest building, the intermediate school, is eight years old. Concern is now centered on the high school, badly in need of repair; for example, because the bleachers are dangerously rotten, one team has requested the best seats be reserved for their supporters when they play Small Town in the fall. But spirit in the district is high. The current administration, led by superintendent Ray Holder, a former coach and high school principal in Small Town, is popular, and recent test scores document continued improvement in student achievement. Also, the district lacks two features of today's public education -- a teachers' union and a district computer -- and neither is missed.

Small Town, South's, Evaluation

Operating now for 250 children in grades K-8 and having lost a third of its staff this year because of funding cuts,

the Title I Program has existed in Small Town since 1965. In December of that year, a social studies teacher and a home economics teacher were called to the superintendent's office, relieved of their classroom duties, and asked to write the district's first Title I proposal. The social studies teacher, now an Assistant Superintendent who jokingly labels himself the district's Title I historian, summed up the evolution of evaluation use in Small Town when he commented, "We have moved from evaluation as a necessary evil to planned evaluation coordinated with other (i.e., regular) programs." This shift has taken place largely since 1977 when the district hired a new administrative assistant for federal programs.

The Small Town, South, Evaluators

Sally Ann Lopenia, an attractive and energetic woman in her late 30's, was well qualified to head Small Town's federal programs. A state native, she taught for eight years in public schools in the south, earned a Masters in Education, and had been the executive director of a non-profit corporation that, at the end of her tenure, supported 22 staff members with funds from both public and private sources. While there, she had also received training at the National Training Laboratory, although she never formally studied program evaluation. Because her husband was a professor at a nearby college, she had lived in Small Town for several years and was committed to staying.

But appropriate as they were, these external qualifications were not the key ones Mrs. Lopenia brought to the job.

A bright woman who learns continually, she made the Title I philosophy that of "taking the soundest educational theories that fit into remedial programs in a way that coincides with district concerns." Recognizing that the Small Town I.S.D. was not interested in "going out on a limb" educationally and that the community held a "no frills" view of good education, she set as her goal the meaningful coordination of Title I instruction with the regular school curriculum. The federal guidelines were a "big stick" she could use for adding innovations like a testing program and a parental advisory group, but her attitude was a positive one, i.e., given that we have to do these things, how can we make them work for us? Knowing that accountability was being forced on the district from outside, she sought to make the experience worthwhile and valuable from the inside. Her earlier work in training had taught her the need for, as well as the means to, effective communication with her staff. This knowledge coupled with her sense of purpose and positive attitude were the critical attributes she brought to the job.

Because the district is small, Mrs. Lopenia's part in Title I includes several roles that would be handled by different people at a larger site. Not only is she the federal program coordinator and grantsman; in an important sense she is also the Title I liaison with the schools and, at least in part, the district's Title I evaluator. The responsibility for this final role is shared with Mrs. Bess Foley, the former home economics teacher who has been active

in Small Town's Title I programs since their inception and who is, as everyone around her is quick to point out, a positive force in the local school community.

A veteran of over 20 years, with a quick smile and self-assured laugh, Mrs. Foley knows the Small Town schools well. She has taught both elementary and high school in Small Town and years ago became its first guidance counsellor, a function she still performs for the junior high school. Like Mrs. Lopenia, she has had no formal training in evaluation, but has been a willing learner on the job. As she puts it, "Whatever the need might be, I meet it," and her mathematical ability effectively complements the verbal skills of Mrs. Lopenia. That the district's Title I evaluation needs are met is clear. Over the course of the five years that Mrs. Foley and Mrs. Lopenia have been responsible for the evaluation, the use of evaluation information in the district has not only grown, but, as was noted above, has been institutionalized into the regular school curriculum. As will be seen in the following sections, Mrs. Lopenia uses both structured and nonstructured means to promote the use of Title I evaluation results in Small Town.

Utilization

Title I Evaluation: Structured Uses

To speak of evaluation in the district is to speak primarily of standardized, functional level testing, i.e., the district-wide administration of the California Achievement

Test (the CAT) every March 1. From her office in the junior high, Mrs. Foley is responsible for the test ordering, distribution, scheduling, and collection. It is also her responsibility to package the tests and put them on a bus to California, where they are machine scored, and where computerized score reports are printed.

Her biggest challenge comes, however, when the results return to Small Town. Knowing that people will be eager to see the scores, she literally locks herself in her office, putting signs on the usually open doors, then one by one rips the print-out sheets and assembles them into a bound volume for each school and into class sets of individual score reports for elementary parents and secondary students. Title I students' scores must be compiled, again by hand, and statistical gain computations worked on the district's only "computer", a programmable calculator purchased with Title I funds. Mrs. Foley's yearly ordeal is finished when she draws charts and graphs to visually show the results and to "get the information in usable form"; she's "good at drawing pictures" and finds them of value in making presentations.

Once the district scores are tallied, Mrs. Foley shares them first with the superintendent and then with principals and teachers, making separate presentations at each campus as the "principal's invited guest." Parents of students in the kindergarten to sixth grades meet with their children's teachers to discuss the results, while older students are personally given copies of their scores. Mrs. Lopenia uses

the Title I computations to prepare a report for the Parents' Advisory Committee and for the June meeting of the Board of Trustees, along with recommendations for the following year's budget allocations. She also prepares a news release for the local paper. All of this sharing of results is, as Mrs. Foley says, "common courtesy", and, because scores have been good, a pleasurable experience. However, although the sharing is thorough, up to this point, it is not highly unusual. The unusual and creative communication of results for Title I children comes several months after the scores are returned, shortly before the re-opening of school for the fall.

The Small Town Skills Charts. Over the summer, the scores are used to select Title I students. Following this, each August since she was put in charge of Small Town's federal programs, Mrs. Lopenia has sponsored an inservice workshop for all Title I teachers and a small number of invited regular classroom teachers, all of whom are paid a small stipend to attend. The purpose of the Federal Programs Workshop, which has decreased in length from five days to one as a result of funding cuts, is informational, and participants unanimously agree that what they learn at the sessions is invaluable to the continued functioning of the Title I program. As one teacher commented, "How else would you all know what you're doing?" Some topics change from year to year (e.g., testing techniques, ways to work with aides, the meaning of a regression model), but standard topics include

a review of the history and purpose of Title I funding, a discussion of the current federal regulations, and, more importantly for this discussion, a review of the previous year's Title I test results.

Besides the overall Normal Curve Equivalent (NCE) gains for all Title I children (ranging from +.61-9.15 in math and +1.02-5.68 in reading for 1980-1981), results are presented for each teacher's new students (K-6) using an objectives management system called the Small Town Reading Skills Chart. Based on what began as a list of Title I reading skills that has since been incorporated into the regular curriculum, this chart has the students' names down the left side and the various objectives across the top. The previous year's CAT results are used to mark each cell either "+" for mastery, "P" for partial mastery, or "-" for non-mastery. A similar sheet is used to plot the math scores, and, in addition, individual charts are prepared for each student to be passed along to the Title I teachers from year to year. Given these charts, teachers can at a glance see which students need what instruction and then prepare and teach accordingly. The test results have been transformed into a highly explicit diagnostic tool for the Title I classroom, and teachers are able to devise what amounts to an individual educational plan for each of their students.

When Title I teachers first meet their students, the required administration of the Brigance Diagnostic Test, an individual inventory of reading and math skill development,

encourages further use of the charted scores. Early in the school year the Title I teachers essentially validate their class's Skills Charts by again testing children and seeing exactly where their strengths and weaknesses lie. This awareness can then be translated into specific, targeted instruction.

A comment on objectives is in order here. The Small Town system is afloat with learning objectives, having endorsed three separate (although necessarily overlapping) sets: (1) those objectives on the Small Town Reading Skills Chart and the less finalized Math Skills Chart; (2) those objectives from the state's minimum competency testing program for grades 3, 5, and 9; and (3) those general objectives measured by the CAT. As mentioned above, the Title I program also uses a fourth set, those on the Brigance Diagnostic Test. Mrs. Foley has "correlated" the three district-wide sets, marking with one color those objectives common to all three sets and with another those common to any two. Where supplementary data are available, the Skills Charts may be accordingly updated. The important factor here, however, is not so much the specific objectives as the process whereby Title I teachers learn before the beginning of the school year the extent to which their students mastered certain learnings the previous year. As one teacher put it, the Skills Chart "certainly makes teaching, I wouldn't say simpler, but at least I have a goal."

The Federal Programs Workshops. The Federal Programs Workshops are planned, in part, by teachers, and their active participation is a key goal for Mrs. Lopenia. Because of this, the Federal Programs Workshop can also lead to a group discussion of how to improve certain scores. The session not only "gives you a goal for the next year", but it "will help you look back at what went wrong". For example, one year Title I students at the junior high level scored poorly on reference skills. At the workshop the following year, the junior high teachers -- both Title I and regular -- planned how they would address this weakness. As a result, during the course of the year, students learned library skills from the librarian, map skills in history classes, and so on, and, by March, the reference skill scores were improved. The Federal Programs Workshop, then, focuses teachers' attention both as individuals and as a group on the previous year's test scores with an eye to the upcoming year's instruction of Title I students. The involvement of selected regular classroom teachers helps to increase the district-wide understanding of Title I and the more general use of the CAT scores. One measure of the effectiveness of the Federal Programs Workshop is that teachers who have not attended one of the workshops now ask to participate.

Less Structured Methods

If the Skills Charts and the Federal Programs Workshop are the formal and structured means Mrs. Lopenia has developed to promote the use of test information, other less structured

activities also facilitate the use of this and other Title I information in Small Town. The "personal factor" has been noted in the literature on evaluation use, and the informal ways that encourage the use of Title I information in Small Town center on Mrs. Lopenia and, to a lesser extent, on Mrs. Foley.

Sally Ann Lopenia is a believer in communication and devotes considerable energy to keeping the Title I staff and parents informed. When asked what information she receives from Mrs. Lopenia, one teacher expressed the general consensus: "Heavens, there's so much information we get from her!" While some of the information communicated during the course of the year is not related to program evaluation much of it is because, in addition to believing in communication, Mrs. Lopenia believes in using data. Four examples are sufficient to document this commitment.

The first two examples are yearly occurrences, the first taking place at the spring meeting of the Parents' Advisory Council. The parents are presented with the new CAT scores, and, although the PAC has no power to make decisions, the parents do complete a checksheet, answering such questions as, "Are we pleased with this Program?" and "What does the Title I Program need to emphasize next year?" In this way they not only see the test results; they use them to suggest programmatic changes and, over time, to gain a better understanding of the Title I Program's strengths and weaknesses.

The second example of use involves the compilation of

the teacher evaluation questionnaires at the Federal Programs Workshop. The questionnaires administered at the conclusion of the sessions contain 33 items to be rated on a scale of 1-7, poor to excellent (e.g., "I have genuinely committed myself to these sessions," and "The leaders stimulated critical and/or creative thinking"). Mrs. Lopenia analyzes the items separately, then calculates overall scores for the questionnaires. The results are finally written up and returned to workshop participants so they have a sense of the group's reaction to the sessions. In the spring the results are also used to help plan the next workshop.

The third example of data use also involves the Federal Programs Workshop. Curious to measure the district level of misinformation about federal programs, Mrs. Lopenia conducted an informal study last year involving Title I and regular teachers. She devised a six-page survey that quizzed respondents on specific information about federal programs, then asked them to mark from whom they got their information, in what form, and what made it useful to them. Although a measurement expert might question the statistical procedures used, the results lent support to Mrs. Lopenia's workshop. "The most accurate information (that which is most free from distortion) occurs as a result of a planned in-service held immediately preceding each school year"; as could be expected, teachers who attended the Federal Programs Workshop had more accurate information than those who did not attend. The best source of accurate information -- again, as could be

expected -- were the federal programs' teachers and supervisor. Based on these data, Mrs. Lopenia recommended continuing the yearly workshops with expanded attendance by regular teachers and parents, and the establishment of a federal programs network to systematically disseminate accurate information about the programs.

Whether or not this will happen is uncertain. As Mrs. Lopenia described it, the study was done primarily to prove a point to herself. Although she showed its results to the superintendent, the primary use made of them was at the last Federal Programs Workshop where she presented the results to the Title I teachers and they, as a group, discussed how they would work to disseminate federal programs information. However, it is too soon to tell if their informal plan is making a difference.

A fourth and final example of data use occurred in 1979-1980 when Mrs. Lopenia was recording attendance figures for the state Title I report, grades 3 and 6. She was shocked to see the high rate of absenteeism for these students and brought this to the attention of the superintendent and Mrs. Foley. They, too, were shocked and set out to see if this pattern held for older students as well. Because it did, over the course of the next year, the district developed a stricter attendance policy for students in grades 7-12 and a plan to emphasize personal counselling with students who are frequently absent. (A great percentage of these students participate in the Title I Program.) This year Mrs. Foley is

working individually with 51 seventh graders, showing them their elementary attendance records and helping them see the consequences of their actions.

As is clear from these examples, Sally Ann Lopenia both uses data herself and regularly encourages others to do so by giving them information and a related task or problem (suggesting changes in the program, planning next year's workshop, developing a dissemination plan, creating a meaningful attendance policy). The form this information takes may be a report (example 1), a written memo (example 2), a meeting (examples 2 and 3), a personal conversation (example 4), or whatever seems appropriate. Before funding cuts forced her to take on local responsibilities in addition to her federal programs work, the location of her office in the intermediate school and her frequent site visits to Title I classrooms, with and without observation forms, truly enabled her to "stay right on top of things." Two people even commented on her demonstrated ability to foresee and prepare for new mandates (e.g., the use of NCE's). Without exception, Small Town's Title I teachers could think of no additional information they might want concerning the Title I Program. The superintendent put it like this: (citing her use of various ways to communicate information) "Sally is good at making Title I things known... She makes things happen." One teacher labeled her a "unifying force" and the "only source of information for our Title (I) Program." The parents and principals also shared this feeling, suggesting

the potent effect of a single person's commitment to the use of information in a school district.

But she is not alone in this commitment, and a second person also encourages use in Small Town. As was mentioned earlier, Mrs. Lopenia shares the responsibilities for the Title I evaluation with Bess Foley, the junior high counsellor and, as she is nicknamed, the "CAT lady." Although her concern is not solely with Title I, Mrs. Foley's role in facilitating the district's use of information should not be underestimated. She does this in three ways. First, she sees to it that the most current test results are available in the counsellor's or principal's office at each school; for both Title I and non-Title I students she is "real opposed to sticking stuff in the cumulative folder" until more recent CAT scores arrive. Second, like Mrs. Lopenia, Mrs. Foley works with teachers, especially in the junior high school, helping them interpret test results and target their instruction accordingly. Her yearly presentations of CAT results identify her as the district testing person, someone to talk to about testing questions, whether or not a teacher's students are served by Title I.

Third, and most important, her role as the junior high counsellor generates the use of test results not just by teachers or even parents, but by the students themselves. Although again this involves not just Title I students, her personal contact with the children who take the tests encourages them to pay attention to the results. Because

she has been in the district for years and physically maintains each student's testing record, she in a sense knows the students even before they reach her in seventh grade. When they do arrive, she lets them know she cares. As she expressed this, "Somebody still has to be in touch with the classroom, still needs to know that that kiddo needs this or that." The importance of having this contact during grades 7 and 8, the transitional years between childhood and young adulthood, is obvious.

If Mrs. Lopenia's use of information works to improve the Title I Program, then Mrs. Foley's works to help students. The combination is clearly effective. One obvious measure of success is the entire district's enthusiasm over testing. What began as a Title I requirement is now an accepted piece of district assessment. From the kindergarten where "children are getting where they love testing", seeing it as a game, to the junior high where students "actually look forward to getting scores" ("At this age, test scores and cars are the most important -- and girls and money"), testing and the use of results are standard and valued procedures.

Observations

Small Town's unofficial Title I historian said he "has witnessed a major change in the district." Whereas earlier mandated evaluation results were sent to the state, then filed away without being used locally, now the test results

are eagerly awaited by teachers, administrators, and students, then used both as a measure of past success and as a guide to future instruction. This is the result of what Mrs. Foley calls a "testing program that plugs into the school program". The Small Town Skills Charts and Federal Programs Workshop guarantees explicit use of these test data in the Title I Program, and, as was seen above, ongoing activities of Sally Lopenia and Bess Foley lead to additional use of Title I information. Use is both immediate and longterm as people's ideas about the Title I Program have changed from negative and suspicious to positive and supportive. At least four factors have contributed to this success: outside support; the district's smallness; the personal factor; and the effective presentation of information.

Outside Financial and Tecnhical Support

Without outside funding, the Federal Programs Workshop could not have been held; Title I funds pay teachers small stipends to attend this and, prior to funding cuts, related in-service sessions during the year. Also, the federally-funded Tecnhical Assistance Center has provided continual technical expertise to Mrs. Lopenia and Mrs. Foley, neither of whom would call herself an evaluator. As Mrs. Foley put it, "I could not have survived, not have done a Model C evaluation without T.A.C....They educated me." This support has ranged from workshops with teachers to discuss testing terminology, meetings with central administrators to interpret test results, moral support in using new statistical

procedures, to the actual use of a T.A.C. computer when the Small Town calculator broke.

District Size

Because the district is small, one person wearing several Title I hats is actually able to plan, monitor, and evaluate the program. As head of federal programs, Mrs. Lopenia is relatively autonomous -- in some sense the reward for her competence -- and has free access to administrators, principals, teachers, and even students. She knows what is happening in Title I classes because she can be there in person, know individuals by name, and share information as needed. Communication is often on an individual and informal basis.

Personal Concern

Both Mrs. Lopenia and Mrs. Foley want people to use the information they gather; the test scores are "data worthy of being used and shared." As was clear in the discussion above, they both therefore encourage use by working with individuals and groups to make it happen, frequently by presenting people a task requiring their use of the information. Mrs. Foley feels strongly that this caring and concern for individuals is what separates evaluation in Small Town from evaluation in larger systems. "We're doing the same kinds of activities as other evaluators, (but) they only see paper and numbers. I still see children." This contact leads to use at both personal and administrative levels.

Presentation of Results

Written reports, memos, slide presentations, conversations, workshops -- all are used on occasion to communicate results. Users are, in a sense, targeted, then presented information in a form appropriate to them personally.

Summary

To summarize, then, four factors -- outside support, district size, personal caring, and effective presentation -- contribute to the use of Title I information in Small Town. Mrs. Lopenia states her approach to evaluation as follows: "We took the evil of evaluation and made it positive." In so doing, she and Mrs. Foley have created a setting for use, and, in a small town in the south where non-use might be expected, such use is instead flourishing.

A final and unfortunate footnote to this description is the uncertain future of Small Town's Title I Program. Despite demonstrated success in the program itself and in the use of its evaluation results, funding was cut this year and, in all likelihood, will again be cut next year. The effects of these cuts on the use of evaluation information will be drastic because if a choice must be made between serving children and making evaluation effective (e.g., with the Federal Programs Workshop), the children necessarily come first. The fact that a computer is the dollar value equivalent of a teacher's aide explains in large part why the district is instead using, as Mrs. Lopenia quipped, "knots

on a counting rope". To the degree that evaluative input has made this program strong, just to that degree will it suffer as evaluation use diminishes. This, coupled with fewer children being served and newer, untrained Title I teachers (as experienced staff migrate to the security of permanently-funded jobs), suggests major challenges for this program in the near future. The sad truth is that such challenges seem unavoidable.

Endnotes

1. The other required data and all process data are generally regarded as "just verbiage". This information is used in the planning process and occasionally, as in the attendance example cited on page 13, in unusual ways.

APPENDIX E

Suburb, West: A Study of Title I Evaluation Use

Richard C. Williams

UCLA

February, 1982

The Setting

The Suburb, West Unified School District (SWUSD) is a relatively affluent community in the West. Its population would rank, in socioeconomic terms, in the upper 10-15 percent of western school districts. The affluence, however, is not evenly distributed. Some enrollment areas are very affluent but have large concentrations of families on welfare. There are numerous single-parent families and the population is becoming increasingly racially heterogeneous. Many of Suburb, West's citizens are engaged in state government at the capitol nearby.

For several years the district was known statewide for its political instability. There were two tumultuous school board recall elections. Recently, however, the district has settled down. The biggest problem the district presently faces is financial. The externally funded budgets continue to shrink and internally funded budgets do not keep up with escalating inflation rates.

The district's instructional program is somewhat more decentralized than most. The district sets parameters around the instructional and curricular programs but each local school has considerable discretion regarding the instructional methods and materials it employs.

The school district's 45,000 students are 90 percent

Anglo, 4.6 percent Hispanic, 2.5 percent Asian, 1.9 percent American Indian, and 1.9 percent Black. In spite of Suburb, West residents' relatively high socioeconomic status, the district's per-pupil expenditure in 1978-79 was $30 below the state average for unified school districts. As might be expected in high socioeconomic districts, the students score well on various norm-referenced tests. For example, the district scored in the 87th percentile in twelfth grade written expression on the State Assessment Program (SAP). Their lowest ranking was at the 68th percentile on the SAP's sixth grade spelling test. Approximately 75 percent of the district's graduates go on to higher education.

The district has designated 18 elementary and two secondary schools as Title I schools. Suburb, West Unified School District is a district with a citizenry that expects to be involved in decisionmaking. Indeed, before they moved to local school advisory committees, the district had a strong communitywide citizen advisory committee. Into that environment, nine years ago, moved Bill Leonard, the evaluation specialist everyone credits with developing Suburb, West's evaluation system. An economist by training, with a good deal of experience in planning techniques, Mr. Leonard has a fundamental belief that evaluations are a waste of time unless they are used by someone. When he

first arrived, the Title I evaluations consisted of sending out computer printouts to the schools. When he visited the schools, his suspicions about nonutilization were soon confirmed and he began the nine-year process of developing the present system.

The Title I program evaluation is conducted by the district's Research and Evaluation Department, which consists of one director and three evaluation specialists. More specifically, almost the entire Title I evaluation responsibility is carried by Bill Leonard and his assistant, Helen Carter.

Several basic premises guide their evaluation program:

- Decisions should be data-informed. Title I advisory committees, who have the assigned responsibility for making Title I allocation decisions, should make their decisions on the basis of data that have been generated about the program. They should learn from previous experience.
- Negative data should not be feared. One learns from failures as well as successes. An atmosphere must be created so that advisory committees will not feel threatened by negative results but will, instead, view negative evaluations as a challenge to correct and improve current practice.
- Title I evaluations should not be used for personnel evaluation. Once teachers see their pupils' test results as a threat to their job security, they may be motivated to take invalid methods to assure high pupil performance (teaching to the test) which ultimately will not provide test data which are accurage measures of pupil performance or program effectiveness.

- Evaluation data do not usually point out specific problems and suggest ready-made solutions. Often evaluations convey mixed messages, e.g., some pupils are making progress and others are not. The implications for action may not always be apparent.

- School administrators, teachers, and parents who serve on Title I advisory committees often do not have group process and decisionmaking skills. They must be given assistance in how to read, analyze, and make decisons upon evaluation data.

In short, Mr. Leonard and Ms. Carter believe that evaluations can be a useful tool in Title I program decisionmaking, but producing evaluations does not automatically result in their being used in program decisions. Utilizing data effectively by Title I advisory committees is almost an unnatural act. Title I evaluation utilization can be greatly increased if a process that specifically addresses the decision needs of local school site advisor committees is developed and implemented. The SWUSD has had this need in mind as it has designed, developed, and implemented an evaluation process almost entirely focused at the local school Title I advisory committee.

Suburb, West's Evaluation Program

The two key Title I evaluation reports in SWUSD are the Academic Performance Profile (APP) and the Educational Program Assessment (EPA). Mr. Leonard has spent nine years developing these two evaluation reports.

The Academic Performance Profile is based on the Iowa Test of Basic Skills (ITBS). The report is derived from data that are reported to the state for Title I compliance purposes. SWUSD, however, has modified its Title I evaluation procedure in two ways to make the Title I assessment data more responsive to district utilization needs. First, Suburb, West has received permission from the state to test only once a year in February. (There is a technical variation in this testing program for first graders in that they are tested in May, but thereafter they are tested on the February schedule.) This February test serves as both a pretest and posttest--that is, last year's test becomes the pretest for this year's February testing. This modification from the standard Fall pretest and Spring posttest serves several purposes. The single administration cuts down on time spent on pupil test-taking. Because the pupils typically have had two teachers in the February-to-February time span, the focus is on program evaluation rather than teacher evaluation. It reduces the confounding effects of fall testing, with pupil skill deterioration over the summer and spring testing, with pupil attention diverted by end-of-the-year activities. It fits into the planning cycle of school Title I advisory councils; that is, the evaluation data can be collected, analyzed, and reported back to the advisory

council for their April-May Title I decisionmaking sessions.

Second, the data are converted from grade level results into Grade Equivalency Scores. The data, which represent approximately 50 pages of computer printout, are reduced to a mere three pages and are presented in three ways:

1. The information is presented graphically in a manner which "tracks" particular groups of students as they progress through the grades (the 1977 first graders as second graders in 1978, as third graders in 1979, etc.).

2. The information is presented graphically in a manner which shows the performance at each grade level for the various years involved (fourth grade in 1978, 1979, 1980, etc.).

3. The profile, in a unique fashion, presents the gains realized from one year to the next for the various grade levels (the gains from second grade in 1977 to third grade in 1978, etc.).

This test reporting is completed on all pupils in all schools in Suburb, West. Test data on Title I participants are reported separately for use by the Title I Advisory Committee. A schoolwide report is generated for use by the School Site Council.

In essence, the district has taken the ITBS data and, instead of simply sending out many pages of computer printout, has developed a reporting format which allows Title I Advisory Committee members to quickly assess the present status of Title I pupils in their schools on ITBS subscales, and to determine the Title I pupils' progress over

time. The data are concisely summarized into graphs and tables which are easily and quickly understood.

The Educational Program Assessment (EPA) is essentially a needs assessment process. It consists of two parts in elementary schools--a parent survey and a staff survey. At the high school level, the parent and staff surveys are augmented by a student survey. The survey has been "exhaustively developed" to cover areas of interest to both state and Title I programs. The basic survey consists of a set of standard questions to which each school can add up to 40 additional questions of its own choosing.

While there are many common items on the various survey forms, there are some differences which address local school needs. The parents are asked about their awareness, concerns, and feelings regarding their children's attitudes and achievements. The staff is asked more specific questions regarding the present status of programs and conditions and they are asked to make suggestions for improvement. These data are summarized into "a highly visual, nontechnical format" which enables rapid interpretation and encourages comparisons. The survey's information is directly usable in the Title I and state program planning processes. It presents the considered opinions of parents regarding the school's program.

The district sees several uses for these data. They

bring to the surface problems that may have been hidden, for example, differences among various respondent subsets. They provide a basis for discussion; they disarm strong, outspoken advocates of a given position who may try to unduly influence an advisory committee to pursue a favorite course of action. Finally, they can be coupled with the APP data to give unique insights. For example, a school staff may feel that the school is doing a good job in teaching reading and the test results may confirm their judgment, yet the parents may view the school's reading program as ineffective. This may point up the need for the school staff to better communicate with parents. Conversely, the parents and teachers may feel the school is doing well in teaching spelling but the test scores may indicate otherwise. This may point up the need to take a close look at the spelling instruction program.

The Educational Program Assessment survey data are collected at different times in different schools; as much as possible each school is allowed to select the most convenient time for the survey. The schools indicate the dates they want the data back and the Research and Evaluation unit guarantees their delivery one week prior to that date.

These two evaluation reports, the APP and the EPA, form the core of the district's Title I evaluation program. They have been carefully developed so that they collect relevant

achievement and attitudinal data and analyze and present the data in a highly visual and easily understood format.

Utilization

Local School Level Uses

By April each school advisory council has all the data needed to begin deliberating its next year's Title I budget: the budget allocation, test scores, staff and parent attitudes about the program. Conversations at the local school level with principals, parent advisory committee members, and teachers confirm that they had indeed been utilizing the Title I evaluations and planning process in their deliberations and Title I decisionmaking. Each Title I school was utilizing the district evaluation forms in order to receive a Title I budget allocation.

However, the participants identified some problems. All these data, even in the highly condensed and visually clear format the district uses, can be overwhelming to an advisory committee. Some, especially parent advisory committee members, have no idea of how to analyze the data, to make decisions upon it, and to assure that they are in compliance with a thicket of federal regulations.

As an aid to struggling advisory committees, Mr. Leonard has developed a planning and decisionmaking process which lists specific steps in sequence, the data that are available

to inform the decisions, and the specific decisions that must be made. Special forms have been developed by the district for reporting decisions so that they can serve both local school district budgeting and administrative needs and satisfy federal reporting requirements.

A major technique in Suburb, West's school site decision-making process is the use of "component committees." These committees are formed around test and attitudinal data that point out strengths or weaknesses of various school components, for instance, student computational skills or student morale. Each "component committee" explores the relevant data and proposes an allocation of resources to help bolster its component in the school. These proposals are then collated and the entire school advisory committee reaches final consensus in the form of specific school site budget requests.

The final school decisions are recorded in the Research and Evaluation Office and in the Budget Office. Expenditures are made on the basis of these reports. SWUSD's Research and Evaluation, after three years of careful planning, has devised a special system which allows schools to change their Title I budget allocation decisions during the ensuing year if they so desire. The schools call a secretary who, using a word processor, can make the program change and determine whether the budget is in balance. This change can be made

without paperwork and takes about 10 minutes to complete.

Mr. Leonard and the other evaluation specialists are constantly in the schools, working with the school advisory committees, helping them to understand the evaluations and to utilize them in decisionmaking. This was a lengthy activity in the beginning because the data-informed method of decision-making was new to both school staffs and parents. Now that most school advisory committees are familiar with the techniques, the Research and Evaluation staff finds this consulting activity less time-consuming.

District Level Evaluation Uses

Since the local school-level evaluation process is functioning satisfactorily in most sites, the SWUSD's Research and Evaluation Department has turned its attention increasingly toward districtwide evaluations.

The main district activity is to summarize data collected in the several Title I schools. The district is beginning to mine the rich lode of data that has been generated in the Title I schools. Chronological analyses are being made across the various schools. From such data, districtwide trends can be assessed, promising practices can be determined, and unsuccessful schools can be identified and singled out for additional attention and resources. While these data are generated districtwide, their use, ultimately,

is to better inform local school site decisionmaking.

In addition, the Research and Evaluation Department has completed a longitudinal study of Title I participants. This study tracks a comparable sample of Title I participants with non-Title I participants over several years. It involved complex statistical analyses for which Suburb, West had to utilize the computer at a major university laboratory. The ultimate finding confirmed that Title I was making the desired impact; the Title I program was making a difference.

Observations

As can be seen, evaluation in Suburb, West is specifically targeted toward utilization. Every effort is made to assure that the data are reported in a clear, easily understood format, that the data are presented in a timely fashion, and that the implications of the data for decisionmaking are clear. Utilization drives Suburb, West's entire Title I evaluation program.

Before turning to utilization impact, it might be useful to reflect on why this district seems to be so "utilization-focused," when many other districts are not. Several respondents offered their perceptions on this.

Local School Level

Principals, parent committee members, and teachers from two schools offered further comment on factors that inhibit

and encourage evaluation use at the local level. At one school there was, at first, a low return of parent surveys; thus the school site committee did not give too much weight to those data. The parent committee members felt the low return rate was because parents didn't really believe that anyone would give their opinions any attention. Once the parent committee member reported that the data were indeed being used the percentage of survey returns increased. Some teachers felt the ITBS results are not very useful for classroom purposes but they are finding the data give insights into the school's educational programs.

All acknowledged that the principal plays a key role in determining whether the planning process works as intended. Two principals, who are knowledgeable about the utilization in several schools other than their present school, said that the evaluations were being utilized differentially in the several schools and agreed that a critical factor is the principal's openness and interest in the program. They felt that the present design of the program makes the evaluation data easy to use and they predict that an increasing number of principals will be using the evaluations as intended.

The two principals interviewed are using the evaluations and are very enthusiastic about them. They report

that they provide them a data base upon which they can make plans and decisions with their advisory committees. The EPA provides useful data to counter vociferous staff or parents who want to see a particular pet project or change enacted. Finally, it has given the parents and staff ownership of the Title I program; they are involved in decision-making and the energies and commitments of everyone in Title I are increased.

Mr. Leonard has observed the following differences over the nine years since the program was begun:

- The schools are more open to parents and community participation. Previously it was rare to see many parents or citizens working in the schools. Now the schools are much more receptive to outsiders who, in turn, feel more welcome and "at home."

- In the past school personnel feared data, believing that somehow it would hurt the school or be used against it. Slowly the schools have learned that data can be used to identify and solve problems. Critical to this development has been a nonpunative attitude on the part of the school board and central administration.

- Public understanding and appreciation of the schools and those who work in them has increased. Parents often viewed teachers and administrators as lazy, incompetent, or both. Through participating on advisory committees and struggling with teachers and administrators to make sense of the evaluation data and translate those data into effective and workable programs, the parents have grown to appreciate the problems teachers face in improving pupil achievement. Schools and teaching have become "demystified."

- Previously the district attempted to solve educational

> problems on a districtwide basis. Now, the individual school sites utilize evaluation data to identify and respond to local school-site problems. The result has been individualized school programs designed to meet local needs.

District Level

The perceptions of the Elementary Schools Title I Director, the Elementary Schools Curriculum Director, and a School Board member on central district Title I evaluation utilization varied, according to their position.

The Elementary Schools Title I Director's main concern is District compliance with federal guidelines. He reads the districtwide reports but doesn't make any decisions based on them. He gives an annual report to the Suburb, West School Board for their information. Some of his Title I monitors read the evaluation reports before they visit schools, and this gives them a quick overview of the school's problems and programs.

The Elementary Schools Curriculum Director has limited funds to improve school performance. If she were to spread these funds out across all the schools they would have limited impact. So she uses the evaluation reports to identify 10 to 12 schools that are having difficulty in achieving district educational expectations. She targets resources (funds, consultants, materials) toward those schools. She uses the evaluation reports to advise new

parents who call seeking the "best" school for their children. She feels one important impact of the Title I evaluation system has been on the principals' management styles. The principals are increasingly using data in decisionmaking, and they seem to be more comfortable in doing so. She stated, "If the state and federal governments were to discontinue funding for this program we would still continue to use it in our district. It has made an important difference."

The School Board member is new to the board but she has for many years been a keen observer of the board's procedures. She has been a member for several years of a school site council, which is essentially identical to the Title I parent advisory committees in operation. Because the Title I evaluations are targeted at the local school level they have had little or no role in school board decisionmaking. She felt, however, that the evaluations may have an indirect effect in that problems are identified and dealt with at the local level and do not reach the board with as great a frequency as they might in districts without such a process. Her hunch was that the district has fewer "festering, unrecognized problems" because they receive attention at the local level.

Conclusion

The Suburb, West Unified School District seems to make

great utilization of its Title I evaluations. The utilization takes place mainly at the local level in the parent advisory committees, with some limited utilization occurring at the central level.

This relatively high utilization has not occurred by chance. It has been carefully composed and orchestrated over a nine-year period by the Research and Evaluation Department, and especially by Bill Leonard. Evaluation data appear to be used in most schools and the intensity of use is increasing as schools become used to this data-based decisionmaking process.

Utilization at the central level is, expectedly, quite low, given the local school site orientation. It is limited essentially to identifying problem schools for resource allocation purposes and in keeping problem identification and solution at the local rather than central level.

The Research and Evaluation Unit is now turning its attention to developing districtwide analyses of the local school reports and plans. These analyses will be designed to give both the district and local schools insights into the progress that is being made, problems that remain, and promising practices.